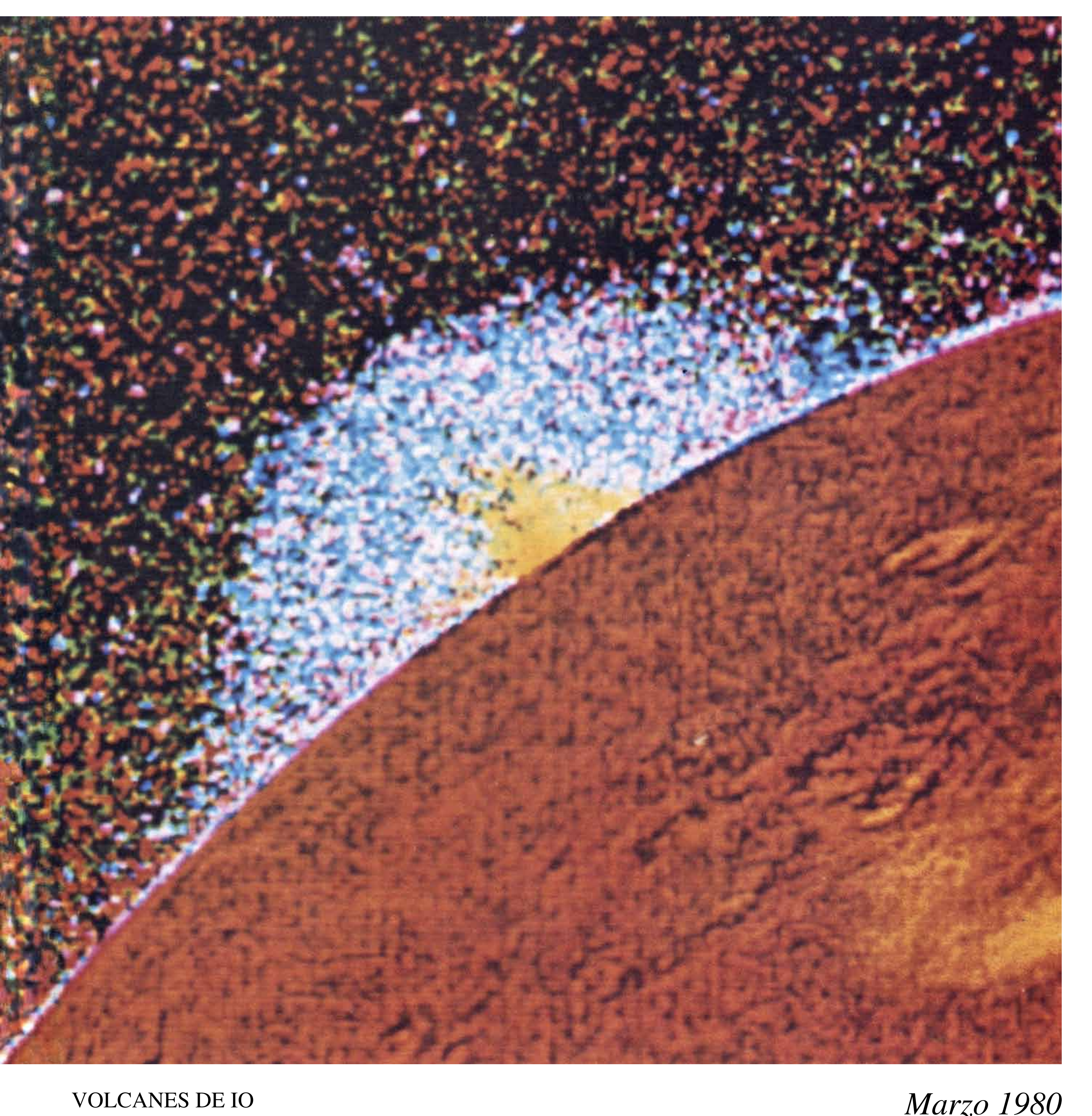
INVESTIGACION Y CIENCIA

Edición en español de SCIENTIFIC AMERICAN



VOLCANES DE IO

Marzo 1980

200 PTAS.

Los espacios en gris
corresponden a publicidad
en la edición impresa

Número 42
Marzo 1980



EDITA Prensa Científica, S.A.
Calabria, 235-239
Barcelona-29 (ESPAÑA)

Colaboradores de este número:

Asesoramiento y traducción:

Ramón Pascual: *La próxima generación de aceleradores de partículas;* Cristóbal Mezquita: *Biología celular del envejecimiento humano;* José M.ª Josa: *Reservas uraníferas mundiales;* Manuel Puigcerver: *Los satélites galileanos de Júpiter;* Luis Bou: *Ilusiones geométricas;* José Luis Blanco: *El origen del maíz;* Antonio Garrido: *La rabia;* Luis Bou: *Juegos matemáticos;* José E. Myro y Alicia Menéndez: *Taller y laboratorio.*

Ciencia y sociedad:

Tomás Jiménez y Antonio Blanco

Libros:

Emiliano Aguirre, Antonio Fontdevila y Luis Alonso



LA PORTADA

La fotografía en la portada muestra el Penacho 2, la segunda de las ocho erupciones volcánicas detectadas en el satélite joviano Io por el vehículo espacial *Voyager I* en el curso de su histórico encuentro con Júpiter el 5 de marzo de 1979. La región simétrica azul, que asciende a unos 200 kilómetros por encima de la superficie del satélite, es la envoltura externa del penacho registrada a través de un filtro ultravioleta. La imagen, reforzada mediante un proceso realizado con ordenador en el Laboratorio de Propulsión a Chorro del Instituto de Tecnología de California, es una "composición ponderada" de cuatro imágenes registradas sucesivamente a través de filtros separados. Las imágenes ultravioleta, violeta y anaranjada se han dividido por la imagen verde y se han usado como azul, verde y rojo en la composición. El penacho, tal como aparece en luz visible solamente, es el que se muestra en la figura de la parte inferior izquierda de la pág. 64 de este número.

Suscripciones:
Prensa Científica, S.A.
Calabria 235-239
Barcelona-29 (España)
Teléfono 322 05 51 ext. 41

Condiciones de suscripción:
España:
Un año (12 números): 2200 pesetas
Extranjero:
Un año (12 números): 36 U.S.$
Ejemplar atrasado ordinario:
225 pesetas
Ejemplar atrasado extraordinario:
340 pesetas

Distribución para España:
Distribuciones de Enlace, S.A.
Ausias March, 49, Barcelona-10

Distribución para los restantes países:
Editorial Labor, S.A.
Calabria, 235-239, Barcelona-29

Publicidad:
Madrid:
Gustavo Martínez Ovin
Avda. de Moratalaz, 137, Madrid-30
Tel. 430 84 81
Cataluña:
Miguel Munill
Balmes, 191, 2.º, 2.ª, Barcelona-6
Tels. 218 44 45 y 218 40 86

Controlado por O.J.D.



PROCEDENCIA DE LAS ILUSTRACIONES

La fotografía de la portada es cortesía del Jet Propulsion Laboratory

Página	Fuente	Página	Fuente
7	Fermi National Accelerator Laboratory		Instituto de Tecnología de California
8-20	George V. Kelvin	70	Ilil Arbel
25	Leonard Hayflick, Children's Hospital Medical Center of Northern California	75-81	Jerome Kuhl
		85-90	Tom Prentiss
26-31	Alan D. Iselin	95	Carey Calloway, Center for Disease Control, U.S. Public Health Service
35	EROS Data Center, National Aeronautics and Space Administration	96-97	Patricia J. Wynne
36	Stephen Deffeyes	98	Martin M. Kaplan y Hilary Koprowski, Wistar Institute of Anatomy and Biology
37-44	Andrew Tomko		
59	Jet Propulsion Laboratory, Instituto de Tecnología de California	100-103	Patricia J. Wynne
		106-119	Eduardo Cadenas
60-61	Ilil Arbel	120-126	Ilil Arbel
62-69	Jet Propulsion Laboratory,	129-133	Michael Goodman

ISSN 0210-136X
Dep. legal: B. 38.999-76
Fotocomposición Tecfa
Guizpúzcoa, 36 (local 1) Barcelona-20
Cayfosa. Santa Perpetua de Moguda
Barcelona
Printed in Spain - Impreso en España

Los autores

ROBERT R. WILSON ("La próxima generación de aceleradores de partículas") ha sido una figura señera en el diseño y construcción de máquinas generadoras de haces de partículas de alta energía a lo largo de más de tres décadas. Nacido en Frontier, Wyo., asistió a la Universidad de California en Berkeley, donde se doctoró en física en 1940. Pasó luego a la Universidad de Princeton; allí colaboró con Enrico Fermi en unas primeras mediciones de las propiedades de absorción de neutrones del uranio 235; en ese período inventó un método de separar isótopos del uranio. En 1943 pasó a Los Alamos Scientific Laboratory, donde dirigió el grupo del ciclotrón y, después, la división de investigación experimental. Fue director asimismo del Laboratory of Nuclear Studies de Cornell University de 1947 a 1967, año en que recibió el nombramiento de director del Fermi National Accelerator Laboratory.

LEONARD HAYFLICK ("Biología celular del envejecimiento humano") trabaja en investigación biológica en el Hospital Pediátrico del Centro Médico de California del Norte. Recibió el grado de doctor en microbiología médica en 1956 en la Universidad de Pennsylvania, donde trabajó varios años antes de integrarse en el Instituto Wistar de Anatomía y Biología. A principios de la década de los 60, mientras investigaba en el Instituto Wistar, aisló e identificó un nuevo microorganismo *(Mycoplasma pneumoniae)* que causaba una neumonía atípica primaria en el hombre. Sus estudios sobre cultivos celulares le llevaron a la conclusión de que las células humanas normales tienen una capacidad limitada para dividirse. Su interpretación de este fenómeno como un signo de envejecimiento celular ha permitido establecer el campo de la citogerontología.

KENNETH S. DEFFEYES e IAN D. MACGREGOR ("Reservas uraníferas mundiales") han trabajado juntos sobre este tema durante varios años. Deffeyes es profesor de geología en la Universidad de Princeton. Se licenció en la Escuela de Minas de Colorado, realizó sus trabajos de postgraduado en Princeton y se doctoró en 1959. Luego trabajó para la Shell Development Corporation como geólogo especializado en petróleo, enseñó en la Universidad de Minnesota y en la Universidad del Estado de Oregon; finalmente, en 1967, volvió a Princeton para incorporarse al claustro de la facultad. A lo largo de los años el trabajo de Deffeyes ha ido desde la geoquímica e hidrometalurgia del sedimentario a los aspectos químicos de la oceanografía y la tectónica de placas. MacGregor es profesor de geología de la Universidad de California en Davis; actualmente no está en Davis pues tiene un contrato de dos años con el Departamento de Energía y está trabajando en Washington. Deffeyes y MacGregor quieren agradecer la ayuda que les prestó James Kukula al realizar los programas de cálculo que se necesitaron para los trabajos a que hacen referencia en el artículo.

LAURENCE A. SODERBLOM ("Los satélites galileanos de Júpiter") es jefe de la rama de estudios astrogeológicos del Servicio de Inspección Geológica de los Estados Unidos. Tiene títulos de Bachelor tanto en Geología como en Física del Instituto de Minas y Tecnología de Nuevo México y un doctorado en Ciencias Planetarias y Geofísica del Instituto de Tecnología de California. Desde que ingresó en la Inspección Geológica en 1970, ha participado en varias de las misiones no tripuladas de exploración planetaria de la Administración Nacional de Aeronáutica y del Espacio (NASA). Soderblom es en la actualidad subdirector del grupo de Experimento Voyager de Ciencia de la Imagen.

BARBARA GILLAM ("Ilusiones geométricas") enseña e investiga sobre percepción óptica y visión binocular en la Facultad de Optometría de la Universidad estatal de Nueva York. Nacida en Australia, estudió psicología en la Universidad de Sydney, y se doctoró con un trabajo sobre visión binocular por la Australian National University. Tras pertenecer algunos años al claustro de la Universidad de Reading, en Inglaterra, emigró a los Estados Unidos en 1967.

GEORGE W. BEADLE ("El origen del maíz") consiguió el tiempo necesario para escribir el resumen de sus investigaciones sobre el origen del maíz (*Zea mays L.*) –un tema de investigación mantenido durante un largo período– al retirarse como Presidente de la Universidad de Chicago, en 1968. Empezó sus trabajos de investigación, hace casi 50 años, como estudiante graduado en la Universidad de Cornell, donde colaboró con el director de su tesis, el fallecido genetista Profesor R. A. Emerson, en determinar "citológica y genéticamente que el teosinte, una planta silvestre, nativa en México y Guatemala, era probablemente el más directo antepasado del maíz". Después de recibir su grado de doctor en genética en 1931, Beadle se fue al Instituto de Tecnología de California, donde continuó sus investigaciones sobre el maíz pero, gradualmente, se fue desplazando hacia el estudio de la mosca del vinagre *Drosophila melanogaster*. Más tarde, siendo profesor de biología en la Universidad de Stanford, colaboró con Edward L. Tatum y otros sobre la relación entre genes y enzimas, siendo su material de experimentación el moho rojo del pan, *Neurospora crassa*. Fue por este trabajo por lo que a Beadle y Tatum les fue concedido, en 1958, el premio Nobel de medicina y fisiología. Beadle volvió al Instituto de Tecnologia de California en 1946, como director del Departamento de biología, puesto que dejó en 1961 para irse a Chicago.

MARTIN M. KAPLAN e HILARY KOPROWSKI ("La rabia") han trabajado juntos, durante casi 30 años, en el estudio sobre la rabia, principalmente en el Instituto Wistar de Anatomía y Biología, donde Kaplan es un asiduo investigador visitante y Koprowski profesor y director. Durante 26 años, Kaplan perteneció a la plantilla de la Organización Mundial de la Salud (OMS), retirándose en 1976; Koprowski, además de su trabajo en el citado Instituto Wistar, enseña investigación médica en la Universidad de Pennsylvania.

EDUARDO CADENAS ("Enzimas alostéricos") es catedrático de bioquímica en la Facultad de Biología de la Universidad de León. Estudió la licenciatura en ciencias químicas en la Universidad de Valencia, donde se graduó en 1954. Obtuvo el doctorado en ciencias en la Universidad de Madrid en 1957. Durante cinco años realizó investigación postdoctoral en los Estados Unidos. Fue Assistant Professor de fisiología en la Facultad de Medicina de la Universidad Vanderbilt y Lecturer de bioquímica y nutrición en la Graduate School of Public Health de la Universidad de Pitsburgh. Después, durante tres años, dirigió el departamento de ciencias fisiológicas de la Facultad de Medicina de la Universidad de Oriente en Venezuela. Con posterioridad ha ejercido la docencia en las Universidades de Madrid, Valencia y Bilbao. Su investigación ha estado centrada sobre sistemas de regulación a nivel hormonal y enzimático.

Hace...

José M.ª López Piñero

... cuatrocientos años

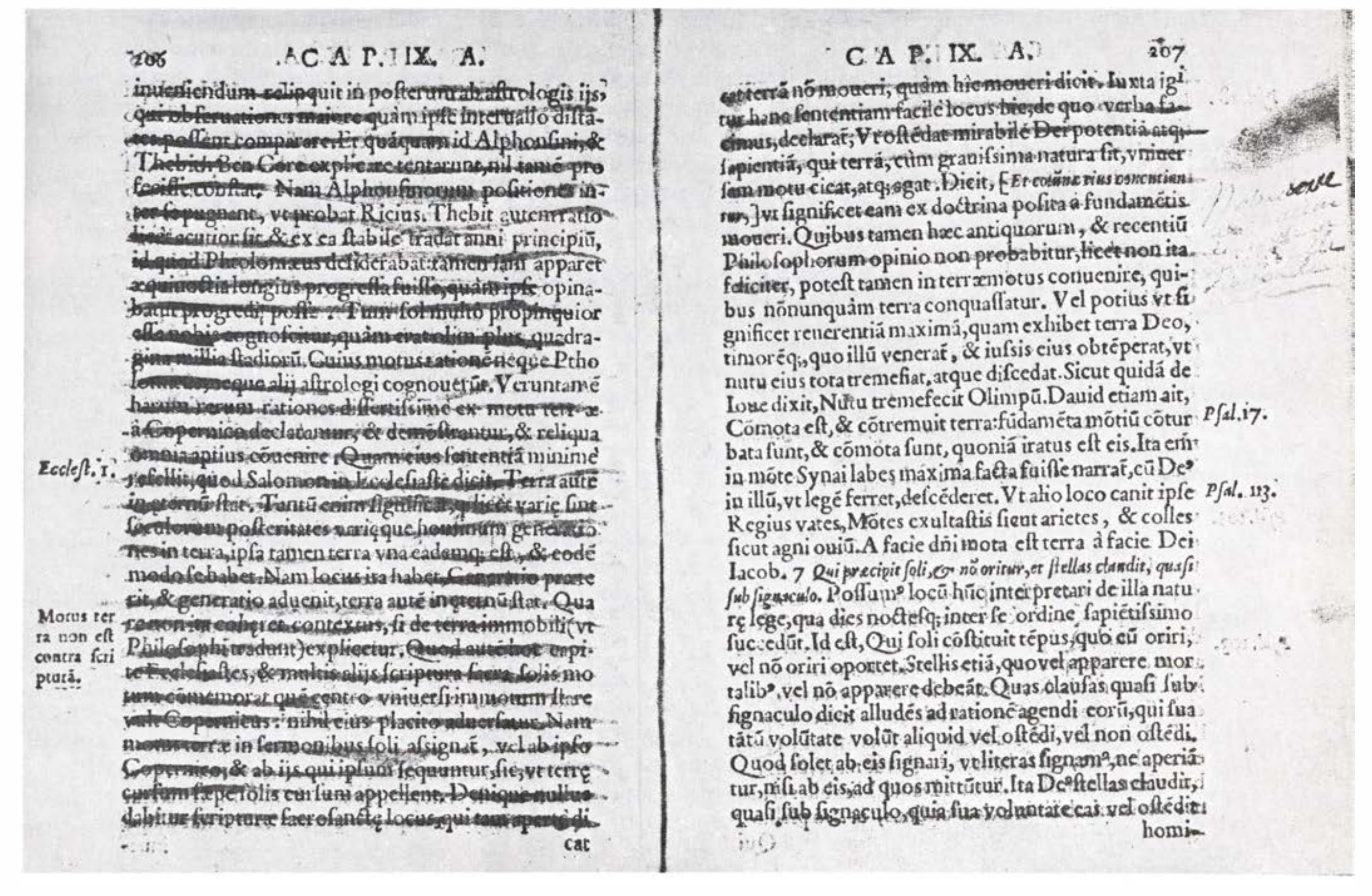
Diego de Zúñiga dejó la cátedra que ocupaba en la Universidad de Osuna y pasó al convento de los agustinos en Toledo, donde residió hasta el final de su vida. Allí escribió su libro *In Iob commentaria* (1584), título de gran relieve en la primera etapa del enfrentamiento con la revolución cosmológica que planteaba la obra de Copérnico.

Al glosar el versículo "Conmueve la Tierra de su lugar y hace temblar sus columnas", Zúñiga defiende en dicho libro dos tesis. La primera de ellas es que el movimiento de la Tierra y, en general, el sistema heliocéntrico de Copérnico, no contradice las Sagradas Escrituras: "Su teoría no contradice en absoluto lo que Salomón afirma en el *Eclesiastés* : 'La Tierra eternamente permanece'. Ello significa que, aunque se suceden en la Tierra varias épocas y varias generaciones de hombres, la Tierra sigue siendo la misma y no cambia. Dice, en efecto, la frase: 'Una generación se va y otra generación viene, mas la Tierra eternamente permanece'. El contexto no resultaría coherente si se hablara de la Tierra inmóvil, como afirman los filósofos. No existe tampoco contradicción porque en este capítulo del *Eclesiastés* y en otros muchos de las Sagradas Escrituras se hable del movimiento del Sol y Copérnico lo considere el centro inmóvil del universo, puesto que en las obras del mismo Copérnico y en las de sus seguidores el movimiento de la Tierra se atribuye al Sol y no hay pasaje alguno en las Sagradas Escrituras que diga tan claramente que la Tierra no se mueve como éste afirma que se mueve".

La segunda tesis de Zúñiga consiste en defender la superioridad del sistema copernicano respecto del tradicional desde el punto de vista astronómico: "Con su teoría, las posiciones de los planetas se explican mucho mejor y de modo más cierto que con la *Magna Compositio* de Ptolomeo y con las opiniones de otros autores. Ptolomeo, en efecto, no podía explicar el corrimiento de los equinoccios, ni presentar un comienzo del año cierto y estable, como él mismo lo reconoce... Por el contrario, Copérnico expone y demuestra de forma muy convincente las explicaciones de estos problemas con el movimiento de la Tierra, concordando todos los demás fenómenos del modo más satisfactorio". A la vista de esta exposición, no resulta extraño que el libro de Zúñiga compartiera con el de Copérnico los honores de la expurgación en el decreto de la Sagrada Congregación del Indice de 1616.

En la España del siglo XVI hubo también, naturalmente, defensores de las tesis opuestas. Por ejemplo, el médico Francisco Valles en su *Sacra Philosophia* (1587), glosa de los temas científicos aludidos en las Sagradas Escrituras. Valles "refuta a Copérnico" con argumentos de la filosofía natural y de la cosmología tradicionales y considera, además, el movimiento de la Tierra incompatible con los textos bíblicos.

Menos rígida fue la crítica de Diego


206 CAP. IX. A.

inueniendum reliquit in posterum ab astrologis iis, qui obseruationes maiore quàm ipse interuallo distantes possent comparare. Et quanquam id Alphonsini, & Thebith Ben Core explicare tentarunt, nil tamé profecisse constat. Nam Alphonsinorum positiones inter se pugnant, vt probat Ricius. Thebit autem ratio [illegible] sit, & ex ea stabile tradat anni principiū, id quod Ptolomaeus desiderabat: tamen iam apparet aequinoctia longius progressa fuisse, quàm ipse opinabatur progredi posse. Tum sol multò propinquior esse nobis cognoscitur, quàm erat olim plus quadraginta millia stadiorū. Cuius motus rationé neque Ptholomaeus, neque alij astrologi cognouerūt. Veruntamé harum rerum rationes disertissime ex motu terrae à Copernico declarantur, & demōstrantur, & reliqua omnia aptius cōuenire. Quam eius sententiā minimè refellit, quod Salomon in Ecclesiaste dicit, Terra autē in aeternū stat. Tantū enim significat, quòd licet varię sint saeculorum posteritates varięque hominum generationes in terra, ipsa tamen terra vna eademq; est, & eodé modo se habet. Nam locus ita habet, Generatio praeterit, & generatio aduenit, terra auté in ęternū stat. Qua ratione non coheret contextus, si de terra immobili (vt Philosophi tradunt) explicetur. Quod autē hoc capite Ecclesiastes, & multis alijs scriptura sacra solis motum cōmemorat, qué centro vniuersi immotum stare vult Copernicus: nihil eius placito aduersatur. Nam motum terrae in sermonibus soli assignat, vel ab ipso Copernico, & ab iis qui ipsum sequuntur, sic, vt terrę cursum saepe solis cursum appellent. Denique nullus dabitur scripturae sacrosanctę locus, qui tam apertè dicat

Eccles. 1.

Motus terra non est contra scripturā.

CAP. IX. A. 207

terrā nō moueri, quàm hic moueri dicit. Iuxta igitur hanc sententiam facilè locus hic, de quo verba facimus, declarat; Vt ostédat mirabilé Dei potentiā atq; sapientiā, qui terrā, cùm grauissima natura sit, vniuersam motu cieat, atq; agat. Dicit, [*Et columnae eius concutiuntur,*] vt significet eam ex doctrina posita à fundamētis moueri. Quibus tamen haec antiquorum, & recentiū Philosophorum opinio non probabitur, licet non ita feliciter, potest tamen in terraemotus conuenire, quibus nōnunquàm terra conquassatur. Vel potius vt significet reuerentiā maximā, quam exhibet terra Deo, timoréq;, quo illū venerať, & iussis eius obtéperat, vt nutu eius tota tremefiat, atque discedat. Sicut quidá de Ioue dixit, Nutu tremefecit Olimpū. Dauid etiam ait, Cōmota est, & cōtremuit terra: fūdaméta mōtiū cōturbata sunt, & cōmota sunt, quoniá iratus est eis. Ita enim in mōte Synai labes maxima facta fuisse narrať, cū Deus in illū, vt legé ferret, descéderet. Vt alio loco canit ipse Regius vates, Mōtes exultastis sicut arietes, & colles sicut agni ouiū. A facie dñi mota est terra à facie Dei Iacob. 7 *Qui praecipit soli, & nō oritur, et stellas claudit, quasi sub signaculo.* Possum⁹ locū hūc interpretari de illa naturę lege, qua dies noctesq; inter se ordine sapiētissimo succedūt. Id est, Qui soli cōstituit tépus, quo eū oriri, vel nō oriri oportet. Stellis etiá, quo vel apparere mortalib⁹, vel nō apparere debeát. Quas clausas quasi sub signaculo dicit alludés ad rationé agendi eorū, qui sua tātū volūtate volūt aliquid vel ostédi, vel non ostédi. Quod solet ab eis signari, vt literas signam⁹, ne aperiātur, nisi ab eis, ad quos mittūtur. Ita De⁹ stellas claudit, quasi sub signaculo, quia sua voluntate eas vel ostédit hominibus

Psal. 17.

Psal. 113.

Páginas de la obra de Diego de Zuñiga **In Iob commentaria** *(1584), con párrafos favorables a la teoría heliocéntrica de Copérnico tachadas por orden de la Inquisición durante el siglo XVII*

Pérez de Mesa que dedicó al problema un estudio, que se conserva manuscrito, titulado *Los movimientos de la Tierra y de los cuerpos celestiales*. Las tesis que defendió en él –reproducidas en sus *Comentarios de sphera* (1596)– eran contrarias al movimiento de la Tierra, pero solamente en términos de mayor o menor posibilidad. Consideraba "posible que la Tierra y los cuerpos celestiales se muevan hacia diversas partes y con desigual velocidad", pero "más posible es que (la Tierra) esté quieta".

Entre ambas posturas existió asimismo una actitud ecléctica, que consideró al sistema ptolomaico y al copernicano como dos teorías igualmente válidas. Dicha actitud, muy cercana al pragmatismo de los astrónomos prácticos, fue también mantenida en una obra doctrinal de carácter escolar como la *Filosofía natural* (ca. 1589) de Pedro Simón Abril: "Ponen todos los astrólogos la Tierra en el centro del mundo, y el Sol en el cuarto cielo, y los demás cielos por su orden, y hacen sus juicios de conjunciones, oposiciones, eclipses y otras cosas semejantes, y salvan las apariencias, y salen los juicios verdaderos. Viene Nicolao Copérnico y trueca la suerte, y hace el Sol centro del mundo, y la Tierra súbela al cuarto cielo y hácela movible, y salva las apariencias, y conforme a este presupuesto hace los mismos juicios y sálenle bien". Como subraya Margherita Morreale, a Simón Abril no se le ocurre siquiera rechazar la teoría de Copérnico por motivos religiosos o filosóficos. Solamente le importa que "satisfacen ambas las dos condiciones esenciales de estar en conformidad con los sentidos y observar las exigencias del razonamiento legítimo".

El movimiento de la Tierra suponía, sin embargo, la crisis de la visión tradicional del mundo físico, contradiciendo –como ha afirmado Koyré– tanto al sentido común como a una filosofía natural muy arraigada. Frente a los habituales acercamientos pragmáticos o superficiales a la obra de Copérnico, el mismo Diego de Zúñiga supo verlo claramente en su *Philosophia* (1597), obra publicada trece años después que sus comentarios al libro de Job. Se debe a Víctor Navarro el haber descubierto la importancia que este texto tiene en la historia del copernicanismo, así como el análisis de las causas que motivaron que Zúñiga acabara considerando imposible el movimiento de la Tierra.

El agustino comienza considerando "absurdísima" la teoría aristotélica del globo ígneo, con argumentos directamente apoyados en Copérnico. Negada la esfera del fuego, el Sol queda como única fuente de luz y calor para la Tierra. En lo que respecta a la situación de esta última "la dificultad es mayor y no se puede tener nada por cierto, por mucho que hombres doctísimos, como Aristóteles y Ptolomeo y otros muchos filósofos y astrónomos intentaran demostrar que el orbe terrestre está colocado en el centro del universo... pues, aunque toda la Tierra diste mucho del centro, la magnitud del cielo puede ser tan grande que para sus habitantes ocurra como si estuviera colocada en el centro del mundo, tal como consta claramente en la gran composición de Copérnico".

El problema fundamental reside, sin embargo, en el "estado de la Tierra". En este punto se produce el cambio de opinión de Zúñiga: "Que la Tierra no está inmóvil, sino que por su naturaleza se mueve, lo opinaron Pitágoras, Filolao... y en nuestra época esto mismo lo enseña Nicolás Copérnico en el libro de las revoluciones y doctamente acomoda la composición del universo con el movimiento múltiple de la Tierra... pero Aristóteles, Ptolomeo y otros filosófos y astrónomos expertísimos mantienen la opinión contraria, a los cuales seguimos nosotros". La causa concreta de la rectificación es el movimiento de rotación terrestre: "Ciertos movimientos que Nicolás Copérnico y otros atribuyen a la Tierra no plantean grandes dificultades. Por el contrario, que toda la Tierra gire en redondo cada veinticuatro horas supone muy graves dificultades y parece que convierte en absurda esta doctrina del movimiento de la Tierra".

Tras superar los obstáculos teológicos y aceptar la superioridad del sistema copernicano desde el punto de vista astronómico, Zúñiga considera insuperables las contradicciones que dicho movimiento implica con la experiencia mecánica basada en el puro sentido común y con los supuestos de la filosofía natural aristotélica. Por ello, aduce argumentos como el siguiente: "Las cosas pesadas que son arrojadas hacia lo alto con fuerza, aunque el tiro sea repetido mil veces, caen de nuevo en el mismo lugar, mientras que si la Tierra se moviera con tanto ímpetu, los sustraería del lugar desde donde fueron lanzados".

Víctor Navarro considera, asistido de plena razón, que esta rectificación resulta ejemplar, porque responde a la conciencia de los poderosos obstáculos filosóficos que se oponían en aquellos tiempos al copernicanismo. Su superación, añade, resultaba "imposible antes de que Galileo fundara la nueva física e hiciera el movimiento de la Tierra *filosóficamente razonable*".

La próxima generación de aceleradores de partículas

Los objetos más pequeños sólo pueden percibirse con los mayores instrumentos. Para seguir profundizando en la estructura de la materia habrá que recurrir a aceleradores cuya futura construcción precisará del concurso plurinacional

Robert R. Wilson

Durante unos 50 años, la investigación de la estructura última de la materia ha avanzado casi exclusivamente por medio de una única técnica experimental. A una partícula de materia se le proporciona una velocidad elevada y se le hace chocar contra otra partícula. A partir del análisis de los residuos liberados después de la colisión se obtiene información sobre la naturaleza de las partículas y sobre las fuerzas que actúan entre ellas. Para llevar a cabo un programa de tales experimentos es necesario disponer de una fuente de partículas energéticas. Los rayos cósmicos constituyen una fuente natural, pero su flujo de partículas es difuso y está fuera del control del experimentador. Una fuente más práctica la constituye un acelerador de partículas, el instrumento que aumenta la velocidad de una partícula y, por tanto, también su energía.

Uno de los primeros aceleradores de partículas, construido por Ernest O. Lawrence en 1928, estaba hecho de cristal de laboratorio de unos cuantos centímetros de diámetro. La mayoría de los aceleradores actualmente en servicio son descendientes directos del instrumento de Lawrence, pero han crecido enormemente en tamaño y complejidad. Los mayores ocupan una extensión de muchos kilómetros cuadrados; sus mastodónticas proporciones han determinado que la posibilidad de su ubicación constituya una condición importante del diseño. El acelerador de partículas ya no es un aparato instalado en un laboratorio; sino que es el laboratorio lo que se construye alrededor del acelerador. La fabricación de semejante máquina "nuclear" cuesta centenares de millones de dólares; y en su funcionamiento interviene una plantilla integrada por unas 1000 personas y docenas de computadores digitales.

Nos hallamos ante el horizonte de una nueva generación de aceleradores de partículas. Los primeros ya han entrado en funcionamiento; siguen otros en fase de construcción; los hay, por fin, en fase de diseño y todavía no están fijadas sus características. Tanto para el físico como para el profano el interés mayor que despiertan esas nuevas máquinas reside en los resultados de los experimentos que posibilitarán. Pero los propios aceleradores también merecen nuestra atención. En la física de las partículas elementales, la energía máxima disponible representa una frontera que marca uno de los límites del conocimiento experimentalmente verificable. Varios de los nuevos aceleradores lograrán energías mayores que las de cualquier máquina existente, y así desplazarán dicha frontera más allá, hacia un territorio no explorado. Para alcanzar esas energías, los aceleradores serán, necesariamente, mayores, más complicados y más caros que sus antecesores.

Debido principalmente al coste, la construcción de un acelerador exige hoy en día la solución, no sólo de problemas técnicos, sino también de problemas políticos, económicos y empresariales. El dinero para la investigación científica es una fuente escasa y es necesario sacarle el máximo rendimiento posible. Las innovaciones técnicas han permitido una reducción sustancial del coste, por unidad de energía, de acelerar una partícula. Es esperanzador observar que se está llegando a otro medio de minimizar el gasto total mundial: a través de la cooperación internacional, puede evitarse la duplicación innecesaria de medios y los proyectos demasiado ambiciosos para una sola nación pueden ser abordados por grupos de naciones afines y quizá, si llegara el caso, a través de una cooperación mundial.

Las grandes inversiones que se hacen actualmente en los instrumentos destinados a la investigación en la física de alta energía sólo pueden justificarse porque las generaciones anteriores de aceleradores han mostrado ya su utilidad. Cincuenta años atrás, sólo se conocían dos clases de partículas aparentemente indivisibles: el electrón y el protón. El tercer constituyente del átomo, el neutrón, se descubrió en 1932. En los años inmediatos se identificaron varias partículas más, mediante experimentos con rayos cósmicos y con los primeros aceleradores. Entre las primeras partículas estaba el positrón, antipartícula del electrón. Aparecieron, asimismo, el neutrino, una partícula sin masa y sin carga eléctrica, y el muon y el pion, dotados de una masa intermedia entre la del electrón y la del protón.

Llegamos así a la década de 1950. En esa época empezaron a funcionar aceleradores más potentes. Y hubo una inesperada, y en algunos aspectos alarmante, proliferación en el número de las partículas conocidas. En pocos años, la lista se extendió a más de 100, la mayoría de ellas clasificadas como hadrones, o partículas nucleares. Entre los hadrones había varios con la nueva propiedad de la materia llamada extrañeza. Hace cinco años fue necesario añadir una nueva clase de hadrones, portadores de otra propiedad de nombre caprichoso, encanto. El ritmo de los descubrimientos sigue aumentando. Se han observado partículas que aparentemente señalan la existencia de dos clases más. Estas clases novísimas, cuya catalogación se ha incoado recientemente, se caracterizan por mostrar propiedades llamadas verdad y belleza o cima y fondo.

Durante un tiempo, pareció que todas esas partículas podían tener un mismo "status" de objetos elementales. Tal posibilidad preocupaba sobremanera, pues resultaba difícil reconciliarla con la con-

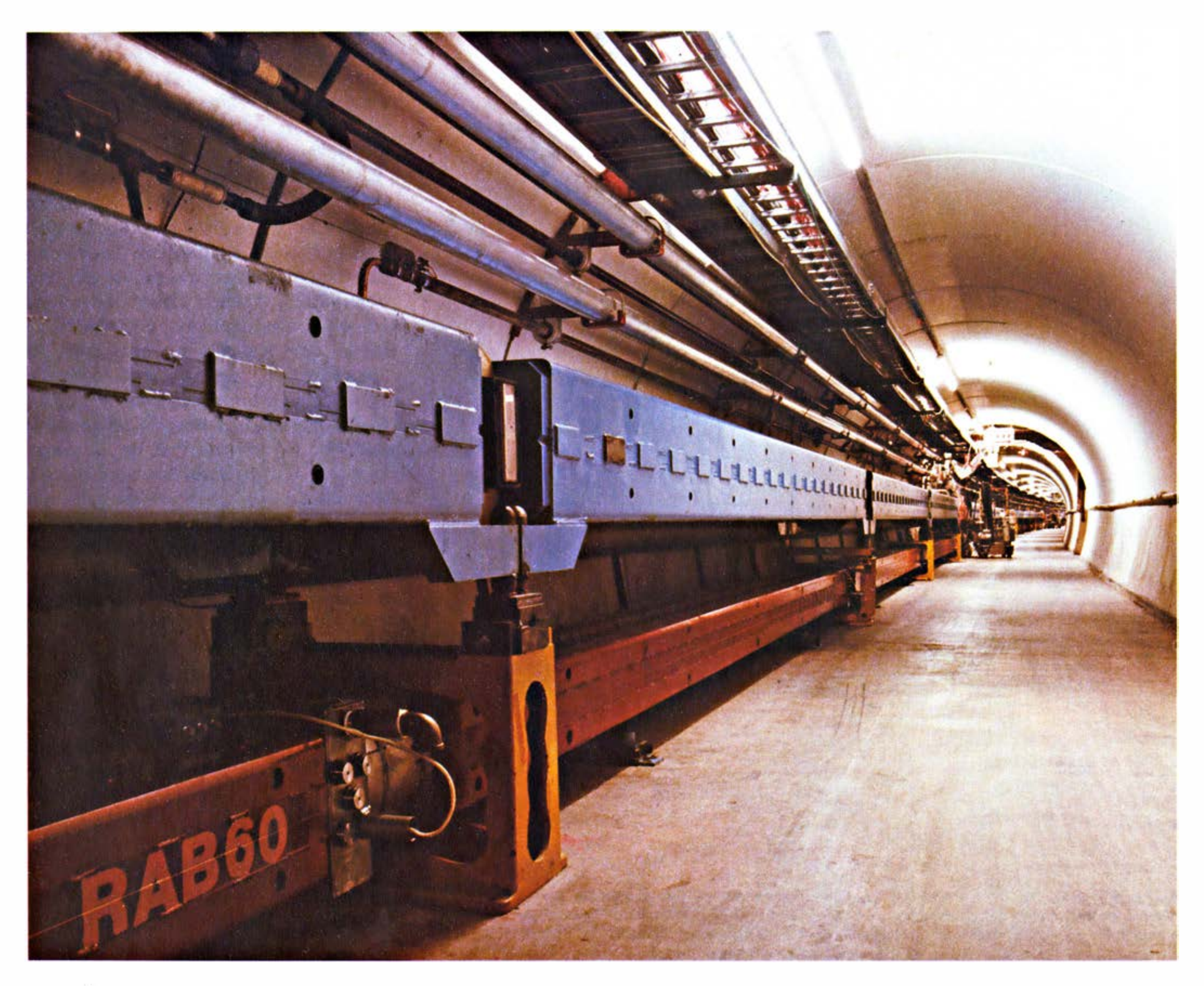
vicción de que las leyes de la naturaleza debían ser razonablemente sencillas. Posteriormente se descubrió, sin embargo, que todos los hadrones podían encuadrarse en esquemas lógicos, algunos de los cuales poseían una bonita forma de cristal de nieve. Además, cabía justificar la existencia de tales esquemas si se suponía que todos los hadrones no eran elementales, sino que estaban constituidos por los entes más fundamentales a los que se dio el nombre de quarks.

Según el punto de vista hoy dominante entre los físicos hay sólo dos clases de partículas elementales: leptones y quarks. De los leptones, la partícula más "popular" es el electrón; comprende también el muon y dos tipos de neutrinos (un neutrino asociado con el electrón y otro con el muon). Hace pocos años se descubrió un nuevo leptón, al que se llamó tau. Es verosímil que el tau posea, asimismo, un neutrino asociado, de manera que debe haber seis leptones en total.

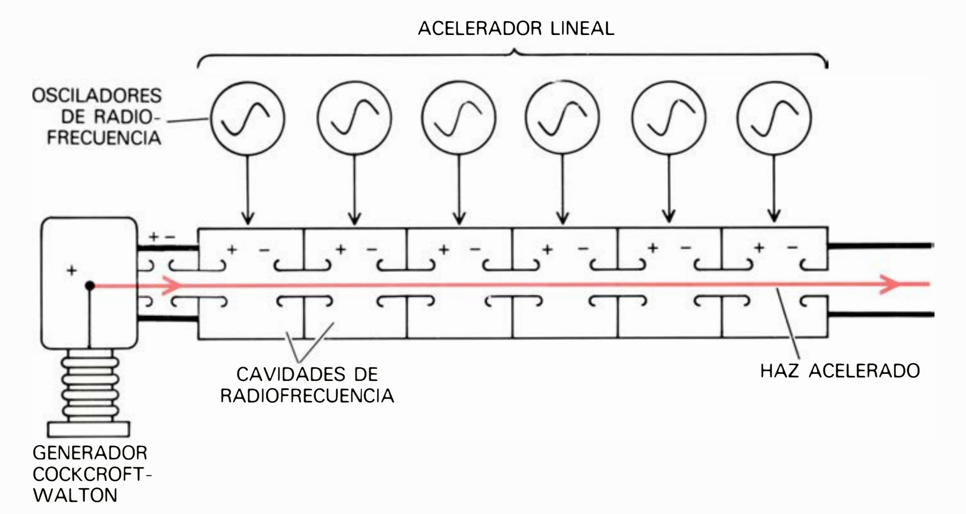
Al parecer, también hay seis clases de quarks a saber: arriba ("up"), abajo ("down"), extraño ("strange"), encantado ("charmed"), cima ("top") y fondo ("bottom"). (No disponemos todavía de prueba experimental del quark cima, pero como todos los demás quarks y leptones se presentan a pares, se supone que el quark fondo tendrá su compañero.) Aunque nadie ha observado un quark aislado, hay razones de peso para creer en su existencia. Todo hadrón conocido (y suman ya algunos centenares) puede explicarse como combinación de quarks o de quarks y antiquarks, formada según reglas explícitas.

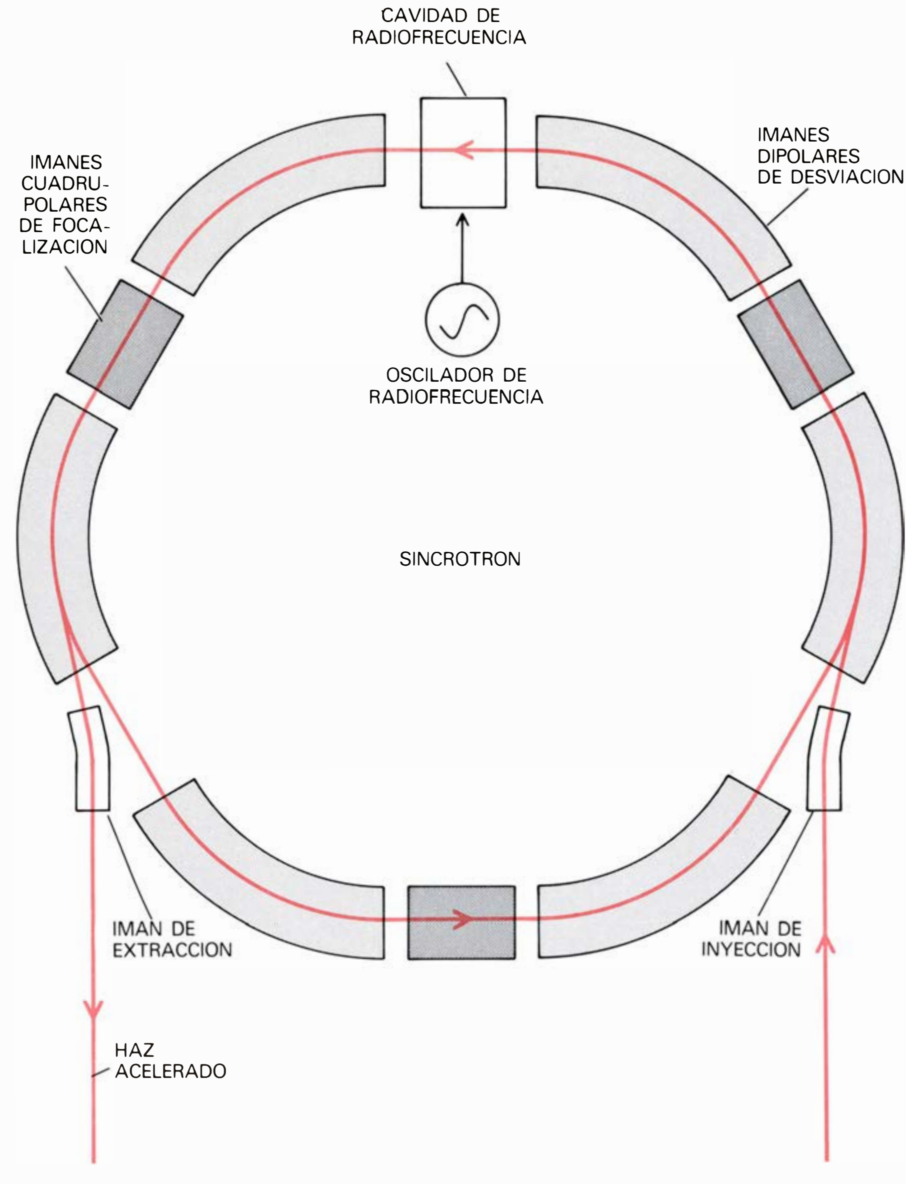
La catalogación de las partículas elementales constituye sólo la mitad del problema global de la comprensión de la estructura de la materia. Hay que conocer asimismo las fuerzas que mantienen las partículas juntas y dan lugar a sus movimientos. Suelen reconocerse cuatro fuerzas básicas, que, en orden de magnitud creciente, son: la gravitación, la fuerza débil (responsable de la desintegración radiactiva beta), el electromagnetismo y la fuerza fuerte. La gravitación y la fuerza débil son universales: actúan entre todas las clases de partículas. El electromagnetismo sólo tiene un influjo directo sobre las partículas dotadas de carga eléctrica. La fuerza fuerte sólo opera sobre los hadrones o sobre sus constituyentes, los quarks.

Se cree que cada fuerza se transmite de un punto a otro por el intercambio de



TANDEM DE ACELERADORES emplazados en un único túnel en el Fermi National Accelerator Laboratory (Fermilab). Los objetos azules de la hilera superior son los imanes desviadores del haz del "anillo principal" original del protón-sincrotrón, que empezó a funcionar en 1972 y ha alcanzado una energía máxima de 500.000 millones de electronvolt (500 GeV). La hilera inferior de imanes rojos es el primer segmento de un nuevo sincrotrón de protones, llamado tevatrón, que está diseñado para alcanzar un billón de electronvolt (TeV). Los imanes del tevatrón tienen arrollamientos superconductores. La escala de los aceleradores puede inferirse a partir de la ligera curvatura del túnel que tiene una circunferencia de 6,3 kilómetros.





ACELERACION DE UNA PARTICULA, como resultado de la fuerza aplicada por un campo eléctrico. Un alto voltaje estático suministra el campo a un generador Cockcroft-Walton, pero la energía máxima que puede alcanzarse por este método es de un millón de electronvolt (MeV). Las energías más altas exigen muchas etapas de aceleración, en las que el campo eléctrico se genera por osciladores de radiofrecuencia. Un acelerador lineal, o linac, dispone de una sucesión de muchas cavidades de radiofrecuencia sincronizadas, de manera que una partícula recibe un pequeño empujón a su paso por cada una de ellas. En un sincrotrón los muchos osciladores de radiofrecuencia pueden sustituirse por una cavidad a través de la cual el haz pasa muchas veces. Un determinado número de imanes desviadores y focalizadores confinan las partículas en una trayectoria que es circular; sin embargo, los campos magnéticos deben ajustarse constantemente a medida que la energía de las partículas involucradas aumenta.

una partícula intermediaria. Para el electromagnetismo, la partícula intermediaria es el fotón: el cuanto de luz y de las otras formas de radiación electromagnética. Así la repulsión entre dos electrones se describe como producida por el intercambio de fotones que son emitidos por un electrón y absorbidos por el otro. (Una repulsión similar se observaría entre dos niños que jugaran, patinando en una pista de hielo, a lanzarse mutuamente una pelota.) El mecanismo de interacción es similar para las restantes fuerzas, excepto en lo que se refiere a la identidad de la partícula intercambiada. El cuanto de la gravitación es el gravitón, y la fuerza fuerte entre los quarks se transmite por las partículas llamadas gluones. En la fuerza débil la partícula mediadora es el bosón vectorial intermediario, también conocido ahora como el "weakón" (de "weak", que significa débil); aparece en tres estados de carga llamados W^+, W^- y Z^0. Mientras que el fotón, el gravitón y presumiblemente los gluones son todos de masa nula, se espera que el weakón sea muy pesado.

Uno de los desarrollos teóricos de mayor relevancia en los últimos años ha sido la demostración de que la fuerza débil y la fuerza electromagnética pueden unificarse. Aunque las dos fuerzas son muy diferentes en sus características observables, pueden ahora entenderse como manifestaciones de un único fenómeno subyacente. La unificación es completamente análoga a la constatación realizada en el siglo XIX de que las fuerzas eléctricas y magnéticas son sólo distintas manifestaciones de la carga. Hoy se está trabajando de cara a la construcción de una "gran unificación", que incluiría la fuerza fuerte, la fuerza unificada débil y la electromagnética. Sin embargo, la fuente última de esa fuerza permanece tan misteriosa hoy en día como las fuentes del Nilo lo estaban en el siglo pasado, y la búsqueda de dicha fuente se halla envuelta en el mismo halo romántico.

Sin embargo, aunque en los últimos 50 años la investigación en el campo de las partículas elementales ha avanzado mucho, sería incorrecto dar a entender que el estado actual de su conocimiento se halle próximo a su total clarificación. El "status" del campo no invita a la tranquilidad. Ni mengua el número de preguntas que piden respuesta. Un primer paquete de asuntos pendientes para los nuevos aceleradores será llenar las vacantes del catálogo de hadrones, particularmente aquellos que incorporan en su estructura los quarks "cima" y "fondo". También importará esclarecer

si la lista de quarks y leptones acaba con los seis de cada grupo que se conocen ahora, o si se encontrarán más a mayores energías. En cierto sentido, seis quarks y seis leptones resultan ya excesivos. Toda la materia ordinaria del universo podría construirse sobre la base de sólo cuatro partículas elementales: el electrón, el neutrino electrónico, el quark arriba y el quark abajo. La existencia de otros leptones y quarks, que aparecen sólo en los experimentos de física de alta energía, es un rompecabezas.

Otro problema lo constituye el fracaso de todos los intentos realizados hasta ahora para detectar un quark libre. A raíz de ese fracaso se han ofrecido varios esquemas teóricos, a fin de explicar por qué los quarks deben estar permanentemente confinados dentro de los hadrones. Existe, sin embargo, la posibilidad de expulsar, de un hadrón, un quark, con tal de que se le suministre energía suficiente. Los programas de futuros experimentos ciertamente incluirán la búsqueda de los quarks.

Una pieza de caza mayor que habrán de cobrarse los nuevos aceleradores será el weakón, transmisor de la fuerza débil. Se estima que las tres clases de weakones tienen una masa aproximada de unas cien veces la masa del protón; el crear partículas tan pesadas excede de las posibilidades de cualquier acelerador existente, y quizá se tarde varios años en conseguirlo.

El desarrollo histórico de la física de las partículas elementales puede considerarse como una extensa lección de escepticismo. En el transcurso de la última centuria, el dominio de la investigación ha progresado, del átomo al núcleo atómico, a los hadrones que constituyen el núcleo y a los quarks que forman los hadrones. Todos estos objetos se consideraron, al menos durante un tiempo, como partículas elementales, sin estructura interna. Quizá sea demasiado pronto para declarar el final de esta progresión, suponiendo que los quarks (y los leptones) sean verdaderamente elementales. Pudieran ser objetos compuestos formados por componentes más sencillos. Los teóricos se han asomado ya al paso siguiente de la secuencia. Así, Haim Harari, del Instituto Weizmann de Ciencias de Israel, ha sugerido que tanto los quarks como los leptones podrían estar formados por sólo dos clases de partículas "últimas". Les llama "rishons", que, en hebreo designa "elemental".

El esquema racional del uso de aceleradores de partículas para explorar la estructura de la materia es claro. Se enfrentan violentamente dos objetos entre sí como puro medio de romperlos en sus partes componentes. Si el choque fuera adecuadamente violento, las partículas podrían reducirse a sus partes últimas e irrompibles. Pero una interacción entre partículas de alta energía no es un equivalente exacto de un accidente de automóvil. No sólo se arrancan partículas del blanco y del proyectil, sino que también pueden crearse otras nuevas a partir de la energía suministrada a la colisión por la partícula acelerada. Por ejemplo, la partícula hypsilón, que se cree incluye un quark fondo, puede formarse en colisiones entre protones energéticos a pesar de que la hypsilón tiene una masa diez veces superior que a masa del protón.

Otra manera de considerar el funcionamiento de un acelerador es por analogía con un microscopio. El límite último de resolución de un microscopio es la longitud de onda de la radiación con que se ilumina la muestra. Los detalles mucho más pequeños que la longitud de onda no pueden resolverse; en consecuencia el microscopio óptico no puede distinguir objetos menores que 10^{-5} centímetros. En mecánica cuántica, la partícula material puede describirse como una onda, que tiene una longitud de onda inversamente proporcional al momento de la partícula. Si nos imaginamos el acelerador como un gran microscopio, al aumentar la energía se reducirá la longitud de onda de la partícula y, por tanto, se mejorará la resolución. El mayor acelerador en funcionamiento tiene una resolución efectiva de unos 10^{-16} centímetros, que es una milésima del diámetro del protón.

Los aceleradores novísimos se sirven de los mismos principios fundamentales que los primeros. La fuerza empleada para acelerar las partículas es el electromagnetismo, por lo que sólo pueden acelerarse las partículas dotadas de carga eléctrica; suele tratarse de protones (con carga $+1$) o electrones (con carga -1). Las partículas se inyectan en una cámara de vacío, que es necesario para evitar que las partículas en movimiento choquen con las moléculas de aire. Un campo eléctrico pone las partículas en movimiento. En el caso más sencillo se aplica un alto voltaje a través de un par de electrodos. Los electrones se encaminan hacia el electrodo positivo; los protones se mueven en dirección opuesta, hacia el electrodo negativo. Un sencillo acelerador de esta clase lo constituye el tubo de la imagen de un receptor de televisión.

La unidad estándar de medida de la energía de una partícula es el electronvolt, que se abrevia eV. Un electronvolt es la energía adquirida por un electrón cuando se acelera por una diferencia de potencial de un volt. Esa misma unidad sirve para medir la energía de protones, o de cualquier otra partícula, en movimiento. Por conveniencia se emplean varios múltiplos del electronvolt para especificar la energía de los aceleradores. El kiloelectronvolt (keV) equivale a 1000 electronvolt, el megaelectronvolt (MeV) a un millón, el gigaelectronvolt (GeV) a mil millones y el teraelectronvolt (TeV) a un billón, o 10^{12} electronvolt. Como la masa y la energía pueden intercambiarse es costumbre dar la masa de una partícula en términos de su energía equivalente, medida en electronvolt. La "masa" del protón es de 938 MeV.

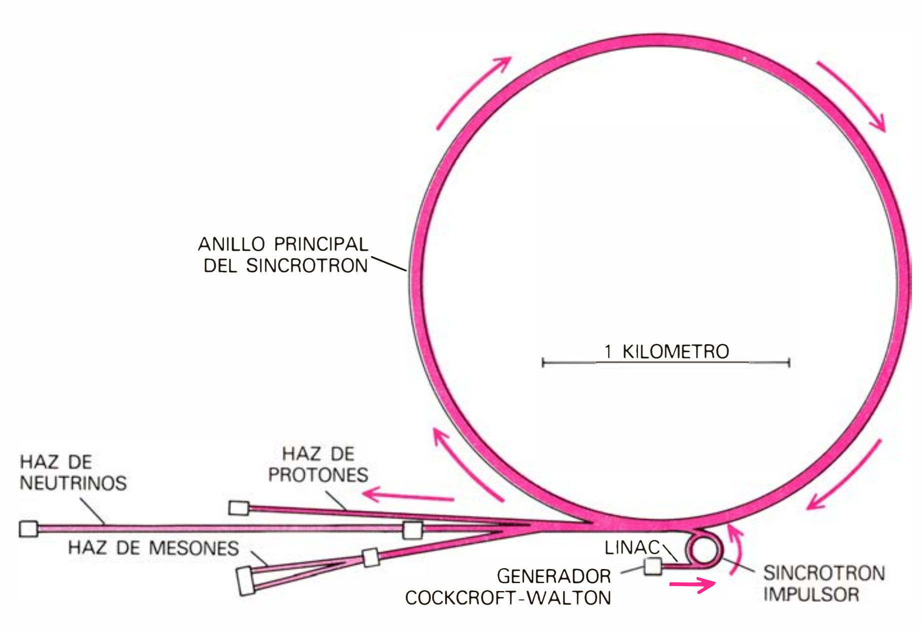
En principio, puede alcanzarse cualquier energía con un sencillo acelerador formado por dos electrodos; basta con aumentar el potencial al nivel apropiado. En la práctica, el máximo potencial que puede mantenerse a través de un par de electrodos es de algunos millones de volt, y así esos aceleradores están limitados a energías de no más de varios MeV. El límite viene impuesto por el inicio de la descarga entre los electrodos o por la rotura de los aisladores.

Si se pretende alcanzar energías más altas, es necesario acelerar una partícula por etapas, dándole una sucesión de pequeños empujes en vez de un empuje grande. La disposición más sencilla de una tal aceleración gradual consiste en formar una hilera de muchas y breves etapas aceleradoras, una detrás de otra. Este es el principio del acelerador lineal o "linac". No se aplica un alto voltaje continuado a los electrodos de cada etapa, sino que se establece un campo eléctrico alternante mediante un oscilador conectado a cada conjunto de electrodos, formando una estructura llamada cavidad de radiofrecuencia. Los osciladores de las cavidades sucesivas se sincronizan de suerte que el campo eléctrico tenga siempre el signo correcto para acelerar, en vez de retardar, la partícula en movimiento. En realidad, una onda electromagnética viaja continuamente a través del tubo de vacío y la partícula cabalga sobre la onda eléctrica como un experto en "surf" cabalga sobre una ola.

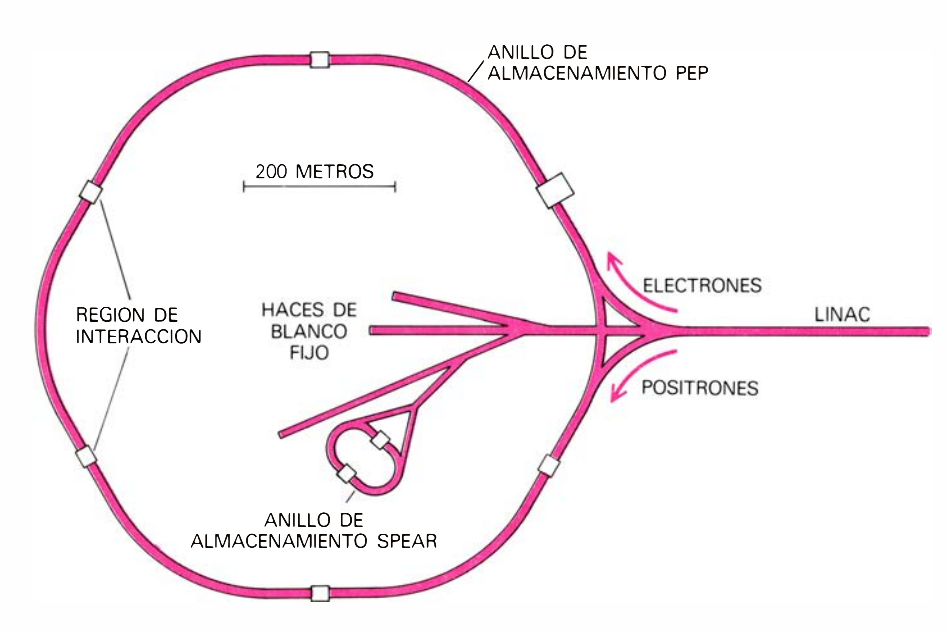
Hay sólo un gran linac en funcionamiento. Se halla instalado en el Stanford Linear Accelerator Center (SLAC), en las cercanías de la Universidad de Stanford. Se acabó de construir en 1961 y costó unos 115 millones de dólares; mide tres kilómetros de largo y está formado por 82.560 cavidades de radiofre-

cuencia. El aparato del SLAC se diseñó para acelerar electrones a una energía de 22 GeV. Actualmente, un programa de sustitución de los osciladores de radiofrecuencia por unidades de más potencia eleva la energía a 30 GeV. Aun cuando no se hubiera mejorado el rendimiento de la máquina, el SLAC hubiera seguido siendo, por algún tiempo más, el acelerador de electrones más potente del mundo.

El límite de la máxima energía práctica de un acelerador lineal lo marca el coste de las miles de cavidades aceleradoras y sus fuentes de potencia de radiofrecuencia asociadas. Para economizar costes, se emplean sólo unas cuantas cavidades haciendo que cada partícula pase a través de ellas muchas veces. Bajo la influencia de un campo magnético una partícula cargada eléctricamente sigue una trayectoria curva. Disponiendo muchos imanes en un anillo, se logra que la partícula siga una órbita circular, o cualquier otra curva cerrada. Un paquete o chorro de partículas puede recorrer el anillo varios millones de veces, pasando a través de las cavidades de radiofrecuencia y ganando energía en cada revolución. Dicho acelerador recibe el nombre de sincrotrón.



ACELERADOR DE BLANCO FIJO, que manda las partículas contra un blanco externo, en el que pueden disponerse detectores y otros aparatos experimentales. El acelerador que se muestra es el del Fermilab. Los protones se aceleran en cuatro etapas: un generador Cockcroft-Walton (0,75 MeV), un linac (200 MeV), un impulsor (8 GeV) y el anillo principal del sincrotrón (de 400 a 500 GeV). Los experimentos pueden realizarse no sólo con los protones acelerados, sino también con haces de partículas secundarias, tales como mesones y neutrinos, arrancados del blanco por el impacto de los protones.



APARATOS DE HACES COLISIONANTES: almacenan partículas de alta energía y las hacen chocar frontalmente. El anillo de almacenamiento que se muestra es el PEP, y se halla en fase de construcción en el Stanford Linear Accelerator Center (SLAC). Los electrones y positrones del acelerador lineal de dos millas del SLAC se inyectarán en el anillo PEP en direcciones opuestas y se mantendrán a una energía de hasta 18 GeV. Las colisiones entre los haces que giran en sentido contrario tendrán lugar en seis zonas de interacción alrededor de la circunferencia donde se podrán construir detectores a lo largo del tubo del haz. El anillo más pequeño, denominado SPEAR, funciona desde 1972 a una energía de centro de masas de hasta 8 GeV (1 GeV es igual a 1000 millones de electronvolt).

Todos los grandes aceleradores de nueva factura que se hallan en fase de proyecto o de construcción son sincrotrones. Vale la pena, por tanto, considerar su funcionamiento con un poco más de detalle. Los imanes que forman el anillo son de dos clases. Los imanes dipolares, que tienen dos polos (un polo norte y un polo sur) generan un campo magnético uniforme; dan lugar a la curvatura de las trayectorias de las partículas. Los imanes cuadrupolares, que originan un campo con dos polos norte y dos polos sur, no desvían las partículas sino que las focalizan en un haz más estrecho, actuando de manera muy parecida a una lente. Repartidas entre los imanes están las cavidades de radiofrecuencia, donde tiene lugar la aceleración. Deben incorporarse también imanes y electrodos especiales para inyectar las partículas en el anillo y para extraerlas del mismo.

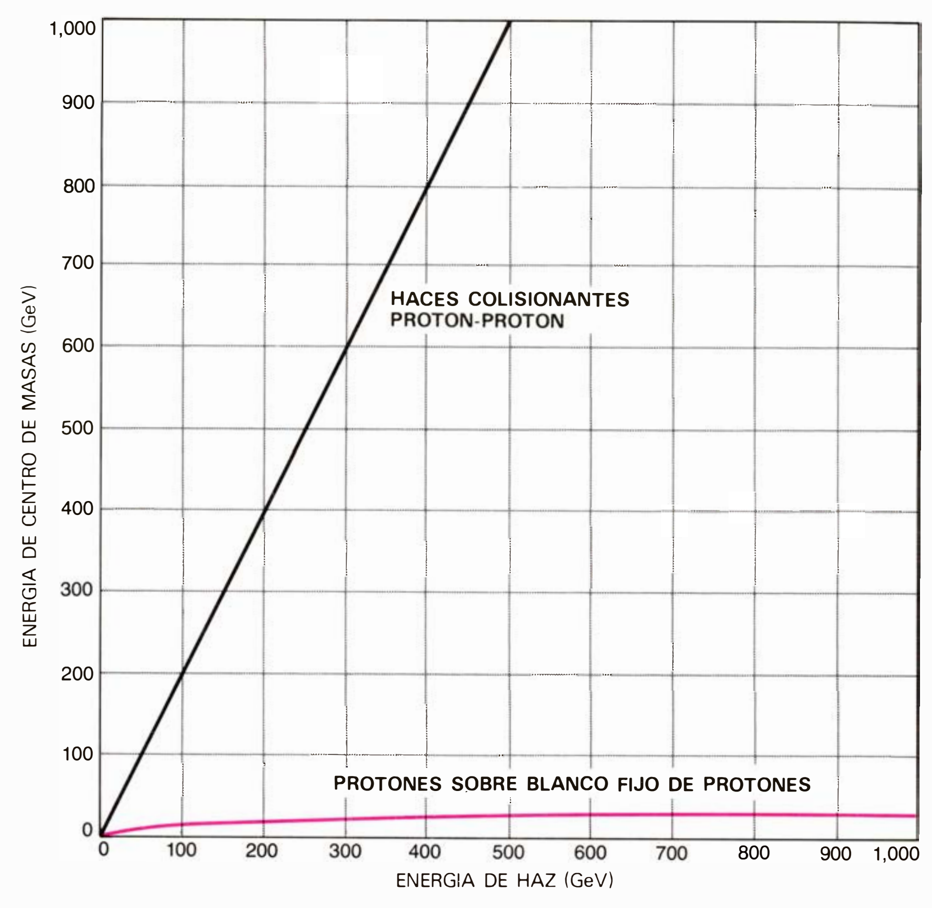
El sincrotrón funciona por ciclos. Se inyecta, en un comienzo, un paquete de partículas, se ajusta el campo de los imanes de desviación de suerte que las partículas sigan, con toda precisión, la curvatura del tubo de vacío. Como la energía de las partículas aumenta en cada revolución, la intensidad del campo de los imanes de desviación debe también aumentarse de un modo suave. Cuando se ha alcanzando la máxima energía, se extrae el haz; se deja entonces que el campo magnético vuelva a su valor original preparándose para el próximo paquete de partículas. Este acelerador recibe el nombre de sincrotrón porque las partículas automáticamente sincronizan su movimiento con el campo magnético creciente y con la frecuencia creciente del potencial acelerador.

Para obtener las mayores energías no basta una única máquina; hay que alinear varias máquinas en serie. Cada una aumenta la energía de la partícula en un factor de 10, e incluso de 100; pasa luego el haz al próximo acelerador de la serie. En varios casos, aceleradores más viejos sirven como inyectores o primeras etapas de máquinas más nuevas y más potentes.

En los aceleradores de protones, la primera etapa suele ser un dispositivo del tipo construido en 1928 por John D.

Cockcroft y Ernest T. S. Walton, del Cavendish Laboratory de la Universidad de Cambridge. Se trata de un gran transformador y rectificador que genera un potencial de cerca de un millón de volt entre un electrodo interno y una capa exterior. Los protones, obtenidos por ionización de átomos de hidrógeno, se depositan en el electrodo interno; cuando emergen (a través de un agujero en la capa), presentan una energía de cerca de 1 MeV.

La siguiente etapa es a menudo un linac, que típicamente eleva la energía del protón a 50 o incluso 200 MeV. Desde el linac los protones se inyectan en un sincrotrón, que puede ser el último eslabón de la cadena o puede servir simplemente para aumentar la energía de los protones e inyectarlos en otro sincrotrón mayor.



ENERGIA EFECTIVA DE COLISION. Esta energía es mucho mayor en los aparatos de haces colisionantes que en los aceleradores de blanco fijo. La energía disponible para crear nuevas partículas es la energía de la colisión cuando la colisión se mira en el sistema de referencia del centro de masas de las partículas colisionantes. En un anillo de almacenamiento la energía en el sistema centro de masas es sencillamente dos veces la energía del haz. (Se supone que ambas partículas tienen la misma masa en reposo y son aceleradas a la misma energía.) Para un acelerador de blanco fijo a baja energía la energía en el sistema del centro de masas es proporcional a la raíz cuadrada de la energía del haz, y a alta energía aumenta incluso más lentamente a causa de efectos relativistas. Los aceleradores de blanco fijo tienen una buena propiedad: mayor luminosidad, que mide el ritmo con que pueden observarse las interacciones entre partículas. Las máquinas de blanco fijo también permiten una variedad de haces secundarios, que no pueden generarse en un aparato de haces colisionantes (*véase figura anterior*).

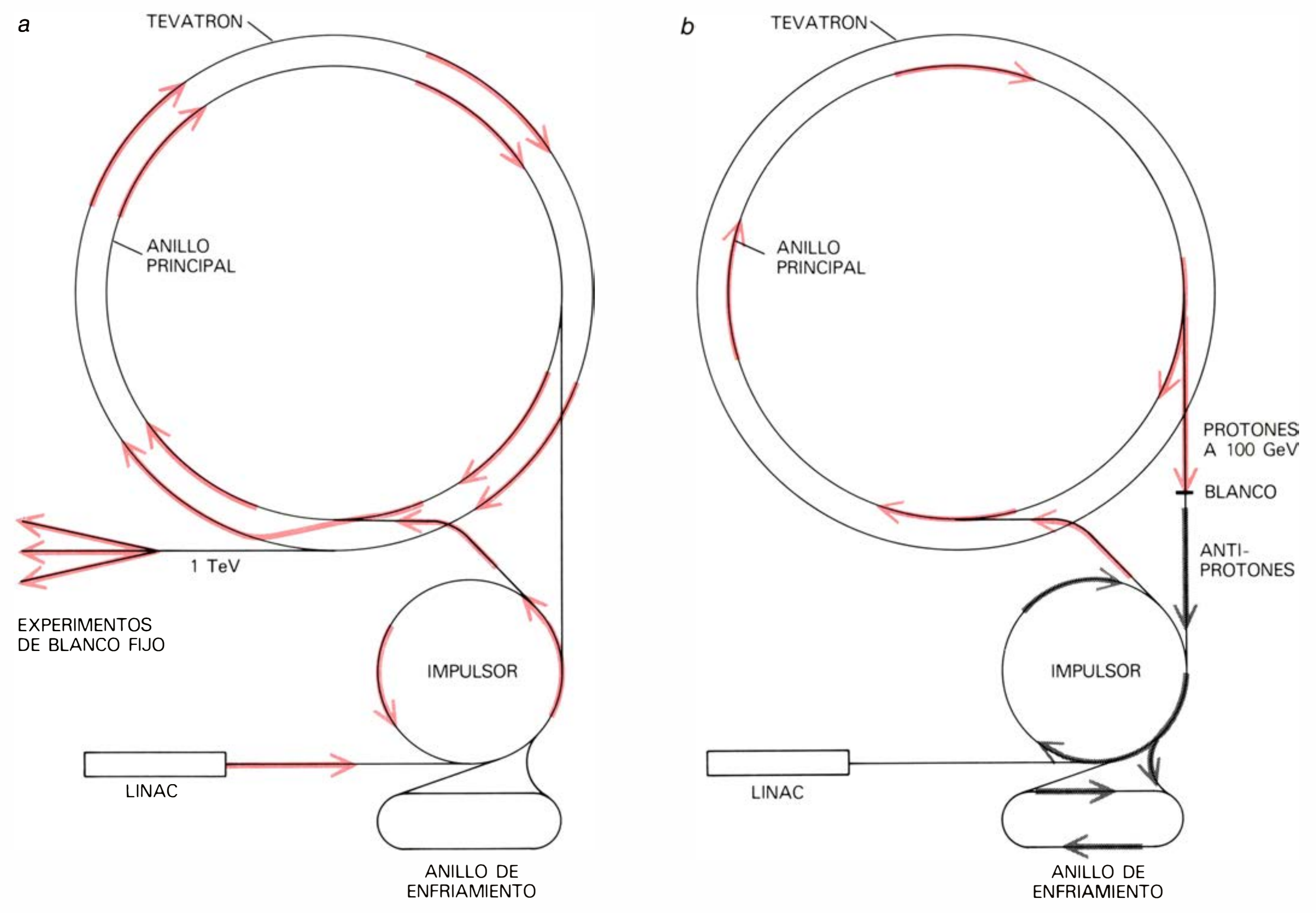
La primera generación de sincrotrones grandes estuvo formada por el cosmotrón, construido en 1952 en el Brookhaven National Laboratory, y el bevatrón, finalizado en 1954 en la Universidad de California en Berkeley. Alcanzaban energías de 3 y 6,2 GeV, respectivamente. En su diseño diferían de las rutinas más recientes, sobre todo, en las referentes a la disposición del campo magnético, que proporcionaba sólo una débil focalización del haz. Al construirse la siguiente generación de sincrotrones, mayores y más potentes, se introdujo un nuevo sistema de fuerte focalización. La forma del campo magnético puede describirse matemáticamente como constituida de una parte uniforme (la componente dipolar) y una parte variable (gradiente) en una dirección perpendicular a la órbita del haz (la componente cuadrupolar). La componente cuadrupolar se hizo más fuerte y su signo alternante, de manera que las oscilaciones de las partículas alrededor de la órbita deseada se hacían más frecuentes, aunque eran de menor amplitud. Como resultado de este gradiente alternante, la separación de los imanes y el espacio para la cámara de vacío podían hacerse menores. A la invención del sincrotrón y de la focalización fuerte debemos que, hoy, el factor económico no sea óbice para la construcción de aceleradores muy grandes.

El principio de focalización fuerte se aplicó por primera vez a los sincrotrones instalados en Brookhaven y en Europa. La máquina de Brookhaven, que se conoce como el Sincrotrón de Gradiente Alternante, abreviado AGS (del inglés, "Alternating Gradient Synchrotron") se terminó en 1961 y alcanzó una energía de 33 GeV. El AGS ha tenido una carrera brillante. A partir de un experimento con el AGS se puso de manifiesto la existencia de dos clases de neutrinos, uno para el electrón y otro para el muon. En 1974, otro experimento realizado en el AGS codescubrió la partícula llamada *J* o psi, que proporcionó la primera prueba del encanto. El AGS sigue en funcionamiento y, como explicaré más adelante, existen planes para un mayor aprovechamiento.

El sincrotrón europeo de focalización fuerte se ha convertido en parte de un esfuerzo mucho más ambicioso, en realidad el mayor de los laboratorios de física de alta energía de todo el mundo. El proyecto fue iniciativa de un consorcio europeo de naciones, que formaron el Consejo Europeo para la Investigación Nuclear, o CERN (del francés "Conseil Européen pour la Recherche Nucléaire"). (El nombre se cambió luego por Organización Europea para la Investigación Nuclear, pero la abreviación ha permanecido.) Constituye un modelo de cooperación internacional, que ha abierto nuevas vías hacia la resolución de los problemas de idiomas, dinero, hábitos de trabajo e interés nacional. Consta hoy de 12 naciones miembros. El laboratorio está situado en la frontera franco-suiza, a unos pocos kilómetros al oeste-noroeste de Ginebra.

El sincrotrón de protones, o PS, del CERN se terminó un año antes que el AGS de Brookhaven. El PS también ha hecho contribuciones sobresalientes. La más importante es que fue el instrumento con el que los físicos del CERN descubrieron un nuevo aspecto de la fuerza débil, llamado la corriente débil neutra, que proporcionó la primera prueba de apoyo a la unificación de las fuerzas débiles y electromagnéticas.

Se sigue investigando con los haces de protones del PS. Aunque la función más destacada de esta máquina, en la actualidad, es, con toda probabilidad, el servicio que presta como inyector para un acelerador mayor todavía, el Supersin-



LA VERSATILIDAD de los múltiples anillos del Fermilab le permitirá funcionar como un acelerador de blanco fijo y como un aparato de haces colisionantes. Los protones incidirán sobre blancos externos (*a*) ya sea desde el anillo principal (a energías de hasta 500 GeV) o desde el tevatrón (a 1 TeV). Un esquema de haces colisionantes emplearía el tevatrón para almacenar protones y antiprotones girando en sentido contrario. Primero se llevarían los protones a unos 100 GeV en ambos sincrotones; se extraerían luego del anillo principal y se emplearían para generar antiprotones (*b*), que

crotrón de Protones, o SPS. En su proyecto inicial, el SPS consistía en un acelerador de 300 GeV, que debía construirse en algún lugar de Europa distinto de la sede del CERN. Sin embargo, en 1965, Francia ofreció un solar junto al laboratorio anterior, sito en territorio suizo. El SPS podía construirse, así, junto al PS. Ni siquiera hubo necesidad de alterar la superficie en la mayor parte del área empleada. El túnel para el SPS, que tiene casi siete kilómetros de circunferencia, se excavó bajo tierra con una máquina perforadora de minería a una profundidad media de 40 metros. El haz llega a la superficie sólo en las zonas experimentales.

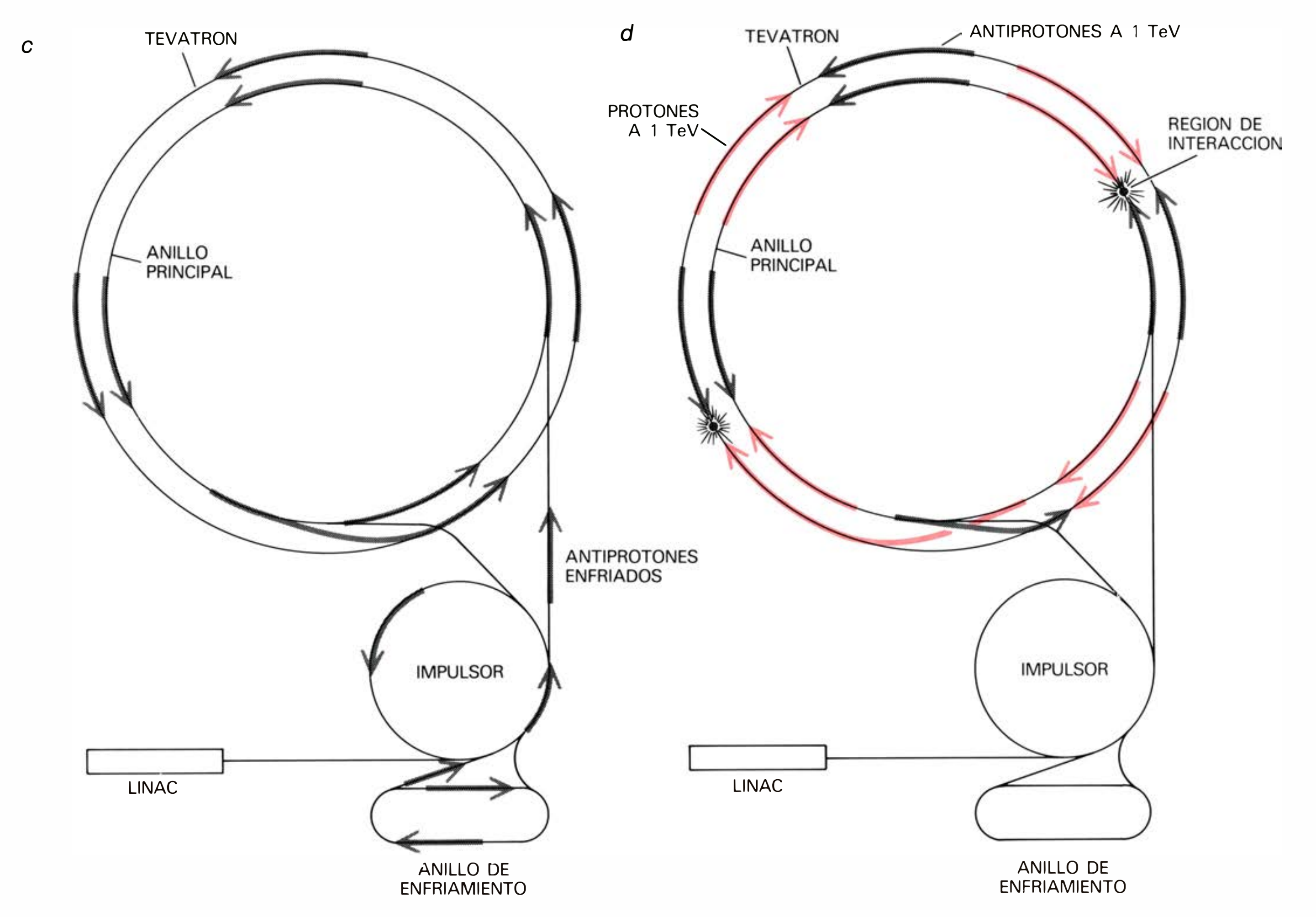
Cuando el SPS empezó a funcionar en 1976, su energía no fue de 300 GeV sino de 400. Tiene capacidad para alcanzar los 500 GeV. No hay duda de que la disponibilidad del PS como inyector aceleró la construcción e hizo factible económicamente la energía más elevada. Los protones de máxima energía del CERN pasan a través de cinco aceleradores: un generador Cockcroft-Walton (550 keV), un linac (50 MeV), un sincrotrón impulsor ("booster") (800 MeV), el PS y el SPS.

Casi al mismo tiempo que empezaron los planes para el SPS se emprendió un proyecto norteamericano de alcance equiparable. Después de un reñido tira y afloja entre estados que querían para sí el nuevo laboratorio, en el año 1967 se eligió un lugar en el campo, a 30 millas al oeste de Chicago. Las instalaciones allí levantadas llevan el nombre de Fermi National Accelerator Laboratory, o Fermilab. No tiene el sello de la colaboración internacional, pero al menos es interacadémico: la autoridad que lo administra es una asociación de 53 universidades. Ni que decir tiene que físicos de muchas naciones experimentan en dicho centro.

Los trabajos del Fermilab avanzaron más de prisa que los del CERN. El acelerador entró en funcionamiento en 1972. El plan original preveía un sincrotrón de protones de 200 GeV, pero se pudo construir una máquina de 400 GeV por el mismo precio. En 1975, la energía máxima había alcanzado los 500 GeV, aunque el coste de la energía eléctrica prohíbe largos períodos de funcionamiento a dicha energía.

El sistema de aceleradores del Fermilab sigue lo que actualmente debe ser un esquema común. Se da un impulso inicial a los protones en un generador Cockcroft-Walton (750 keV), y pasan luego a través de un linac (200 MeV) y un sincrotrón impulsor (8 GeV) antes de introducirse en el anillo principal del sincrotrón. El anillo principal es un poco más pequeño que el del SPS; su circunferencia mide 6,3 kilómetros. Consta de 774 imanes desviadores, cada uno de 6 metros de largo, y 180 imanes de focalización más pequeños. Una serie de cavidades de radiofrecuencia añaden 2,8 MeV a la energía del protón en cada revolución. Pasar de 8 a 400 GeV exige, por tanto, 140.000 vueltas, que se realizan en unos tres segundos. El control de los sucesos rápidamente variables en una máquina tan grande y complicada requiere un conjunto de tres grandes computadores digitales y muchos computadores satélites más pequeños.

El proceso por el cual un acelerador



tienen la misma masa pero carga eléctrica opuesta. Los antiprotones pasarían a través de un impulsor ("booster") a un anillo de "enfriamiento", donde se les convertiría en un haz bien colimado. Cuando se hubiesen acumulado suficientes antiprotones podrían volver, a través del impulsor, al tevatrón, en donde se inyectarían en sentido contrario al de las agujas del reloj, al revés de su sentido usual de rotación (*c*). Los protones y antiprotones se acelerarían simultáneamente hasta 1 TeV, dando una energía en el sistema del centro de masas de 2 TeV. El Fermilab se encuentra en Chicago.

"prima ballerina" se convierte en un simple comparsa de un coro de ballet, sirviendo como inyector de un instrumento mayor, está a punto de repetirse en el Fermilab. Se está instalando, bajo el anillo principal preexistente, un nuevo anillo de imanes y cavidades de radiofrecuencia: habrá, pues, dos aceleradores en un túnel. Los protones se llevarán a una energía de unos 150 GeV en el anillo principal, y se mandarán luego al anillo inferior para posterior aceleración. El nuevo anillo está formado por imanes superconductores, que pueden alcanzar una intensidad de campo doble que la de los viejos imanes. La energía máxima del acelerador se dobla, por tanto, y llega a 1 TeV. En reconocimiento de este récord, se ha bautizado al nuevo acelerador con el nombre de tevatrón.

El uso de imanes superconductores se tuvo ya en cuenta cuando se diseñaron los anteproyectos del Fermilab, pero el riesgo de una tecnología aún no comprobada parecía entonces demasiado elevado. La construcción del tevatrón ha resultado ser un trabajo desafiante incluso para nuestros días. El problema principal reside en enfriar casi 1000 imanes, repartidos en 6,3 kilómetros de túnel, a 4,5 grados Kelvin, la temperatura a la que los conductores especiales de los arrollamientos de los imanes pierden toda la resistencia al paso de la electricidad. Para mantener semejante temperatura, se bombeará un río de helio líquido a lo largo del anillo. Veinticuatro pequeñas plantas refrigeradoras se repartirán a lo largo del túnel; el licuador central de helio será el mayor del mundo, con una capacidad de 4000 litros por hora.

Los arrollamientos de los imanes están formados por una aleación de niobio y titanio inmersa en una matriz de cobre. Se necesitarán unos 30.000 kilómetros de este cable para completar el anillo; ello permitirá calificar al Fermilab con otro superlativo: será el mayor consumidor del mundo de materiales superconductores. A la intensidad de campo máxima los superconductores llevarán una corriente de 4600 ampère, y cuando un imán se "apague", o pierda su propiedad superconductora, la energía almacenada en el campo (cerca de medio millón de joule por imán) deberá disiparse sin destruir los arrollamientos.

Un problema particularmente gravoso ha sido la necesidad de mantener la uniformidad del campo magnético con una precisión mayor que una parte en 1000. Por hallarse los arrollamientos inmersos en su propio campo, están sujetos a una fuerza reactiva de cerca de 400 kilogramos por centímetro lineal. Pero no puede permitirse que los filamentos se muevan ni siquiera dos milésimas de centímetro, ya que ese desplazamiento distorsionaría el campo y podría originar un excesivo calor de fricción. Los arrollamientos se inmovilizan mediante collares laminados de acero inoxidable. La alineación de los imanes se complica también por la contracción térmica que aparece cuando el anillo se enfría a su temperatura de trabajo; un imán de seis metros de largo se contrae en unos dos centímetros.

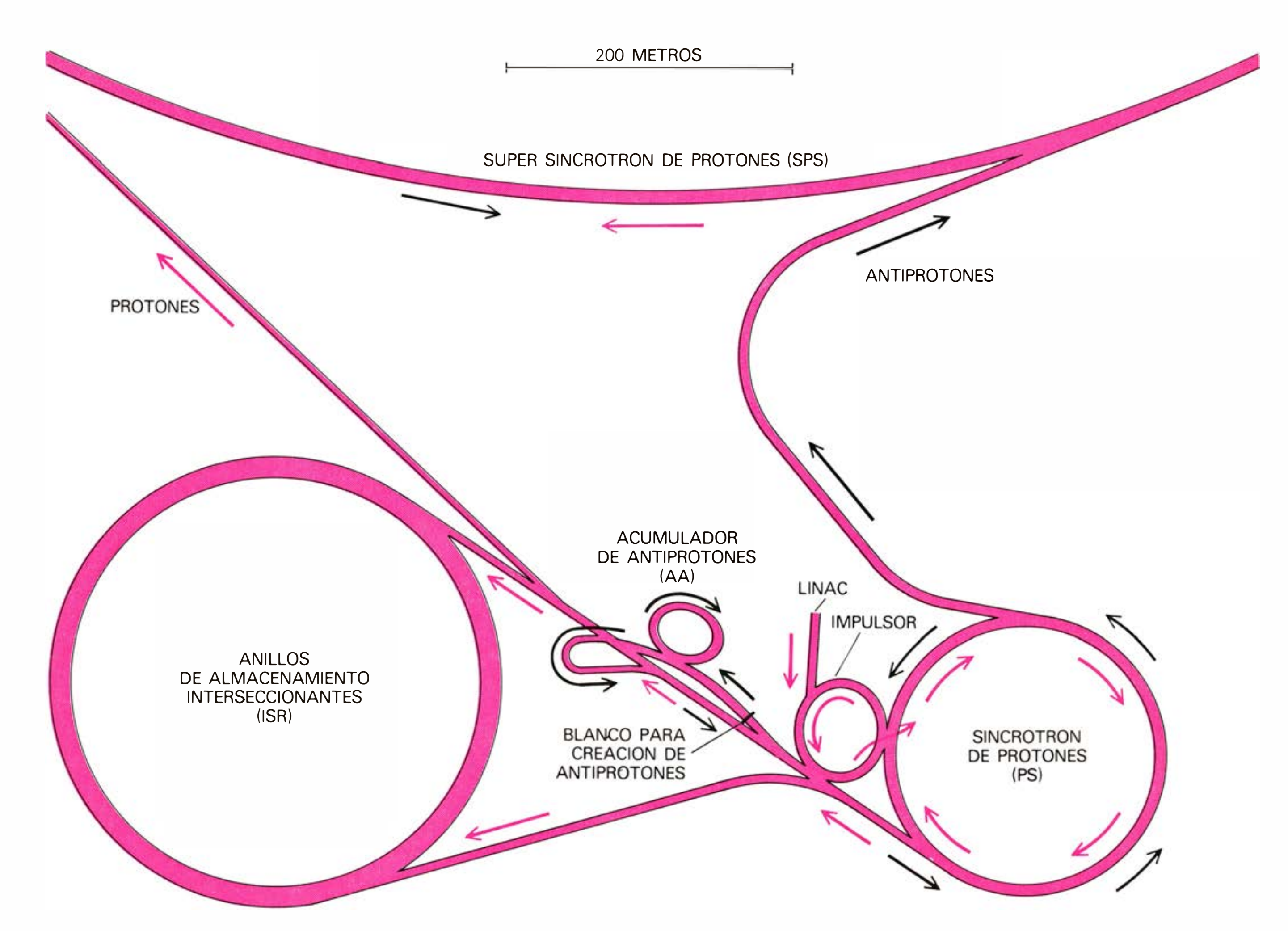
Los físicos tienen la seguridad de que todos esos problemas han sido comprendidos y resueltos. Se ha instalado una cadena de 24 imanes superconductores

en un segmento del túnel y por ella se ha hecho circular un haz, que se ha desviado del anillo principal, a una energía de 90 GeV. Los imanes se comportaron como era de esperar. Se confía que el anillo inferior estará acabado hacia finales de 1981, y los protones, acelerados a una energía de 1 TeV, llegarán a las zonas experimentales en 1982. El ritmo de construcción depende esencialmente del ritmo de financiamiento.

Merece la pena pararse a considerar cuánta energía representa un TeV por protón. En unidades más corrientemente aplicadas a un cuerpo macroscópico, una partícula de un TeV tiene una energía de 1,6 erg, que viene a ser la energía cinética de u. mosquito en vuelo. A plena intensidad, el tevatrón acelerará 5.10^{13} protones cada vez, lo que dará al conjunto del haz una energía de ocho millones de joule. Para hacernos una idea, la energía de una pieza de artillería de unos 45 kilogramos. Si llegara a ocurrir que el haz quedara fuera de control, fundiría las paredes de la cámara de vacío y destruiría los imanes circundantes; obviamente debe evitarse un tal accidente.

Cuando las partículas en un sincrotrón han alcanzado toda su energía, salen impulsadas fuera de su órbita por un imán especial y se desvían hacia una línea de haz exterior. Acabarán por chocar contra un blanco. Las interacciones de los protones con el blanco pueden estudiarse directamente; también es posible crear haces de partículas secundarias arrancadas del blanco. En el Fermilab, por ejemplo, hay zonas separadas para experimentos con protones (las partículas primarias), con mesones (partículas de masa intermedia, tales como el pion), con neutrinos y muones y con fotones. Los experimentos con neutrinos han resultado excepcionalmente gratificadores en los últimos años, porque el neutrino está sujeto solamente a la fuerza débil; por consiguiente, las propiedades de esta fuerza pueden observarse sin interferencia de otras clases de sucesos.

Los detectores y demas aparatos necesarios para los experimentos se construyen a menudo a gran escala, comparable a la de los aceleradores. Para proteger un detector del Fermilab se ha echado mano de la armadura metálica de dos barcos de guerra desguazados. Un detector instalado en el CERN emplea muchas toneladas de fino mármol de Carrara. Estos detectores son "contadores": miden la energía, dirección y poder de penetración de varias partículas y almacenan la información electrónicamente para un análisis posterior. Otra clase de detector es la cámara de burbujas, donde el paso de una partícula cargada eléctricamente deja un rastro indicador de pequeñas burbujas. A partir de las fotografías de las trazas puede reconstruirse la serie de sucesos que siguen a la colisión de una partícula. El Fermilab tiene una cámara de burbujas de 4 metros y medio de diámetro; el CERN tiene tres cámaras, la mayor de las cuales mide unos tres metros y medio de diámetro.



MODIFICACION DE UN SINCROTRON, proyectada, asimismo, en la Organización Europea para la Investigación Nuclear (CERN). Un aparato de haces colisionantes protón-protón, los Anillos de Almacenamiento Interseccionantes (ISR), ha estado funcionando en el CERN desde 1971. Consta de dos anillos entrelazados, con ocho zonas de cruce donde se pueden observar colisiones a energías de centro de masas de hasta 62 GeV. Un proyecto que se está realizando empleará el anillo mucho mayor del Super Sincrotrón de Protones (SPS) para colisiones protón-antiprotón. La fuente de protones para todos los aparatos del CERN es el Sincrotrón de Protones (PS), acabado en 1959. Los antiprotones se crearán como partículas secundarias del haz del PS y se recogerán en un anillo de enfriamiento llamado Acumulador de Antiprotones (AA). Los protones y antiprotones se inyectarán entonces en direcciones opuestas en el SPS y se llevarán ambos hasta 270 GeV. Por tanto, la energía de centro de masas máxima será de 540 GeV.

Las máquinas del CERN y del Fermilab son aceleradores de protones, pero los electrones también pueden servir como material de trabajo de un acelerador. Los principios físicos son exactamente los mismos, aunque las características de la máquina varíen un tanto. La principal razón de diferencia reside en la mayor importancia que tiene, en los aceleradores de electrones, la radiación electromagnética que disipa la energía de una partícula acelerada.

Se sabe desde hace más de cien años que toda carga eléctrica acelerada debe radiar ondas electromagnéticas. En este contexto, la "aceleración" no sólo se refiere a un cambio en la rapidez de una partícula sino también a cualquier cambio en su dirección. Si se le obliga a seguir una trayectoria circular, sufre una aceleración continua, aun cuando la celeridad (o módulo de la velocidad) sea constante. La radiación emitida bajo estas condiciones se llama radiación de sincrotrón, porque se observó por primera vez en el sincrotrón.

La radiación de sincrotrón aparece como un haz intenso de ondas electromagnéticas de alta energía con una distribución continua de longitudes de onda, que se extiende hasta las regiones ultravioleta y de rayos X del espectro. La energía consumida por este proceso debe compensarse mediante un suministro adicional de potencia de radiofrecuencia. Así, la radiación de sincrotrón actúa como una fuerza resistiva análoga al rozamiento.

La energía emitida en forma de radiación de sincrotrón varía en relación inversa a la masa de la partícula elevada a la cuarta potencia. Como la masa del protón es 1836 veces mayor que la del electrón, el problema de suministrar la energía perdida se magnifica para los electrones en un factor de 10^{13}. A las energías de haz que se han logrado hasta ahora, la radiación de sincrotrón no es ningún factor significativo en el diseño de aceleradores de protones, mientras que sí constituye la principal restricción para la energía de los aceleradores de electrones.

Una solución al problema de la radiación de sincrotrón es construir un acelerador lineal donde no haya curvatura y la radiación resultante de la aceleración más suave de los electrones en línea recta sea despreciable. Por esta razón se eligió un linac para el acelerador de electrones del SLAC. De nuevo el principal obstáculo de este método es el coste de las muchas cavidades de radiofrecuencia y de los múltiples osciladores necesarios para excitarlas. En un linac un electrón pasa por cada cavidad sólo una vez, lo que representa un gran derroche. El SLAC valió 115 millones de dólares, el Fermilab casi 250 millones de dólares; pero en éste el coste por electronvolt es más de 10 veces menor que en el SLAC.

Una solución alternativa al linac es construir un sincrotrón de electrones y aceptar el coste de la radiación de sincrotrón. Se puede alcanzar un equilibrio entre los costes de construcción y el gasto continuo implicado en el funcionamiento de los suministros de potencia de radiofrecuencia. La energía radiada por vuelta es inversamente proporcional al radio de curvatura de la trayectoria de la partícula; la pérdida de energía disminuye cuando se fabrica un acelerador mayor. En 1965 se construyó un sincrotrón de electrones de 10 GeV en la Universidad de Cornell, en un túnel excavado bajo una pista de atletismo. En aquella fecha, las proporciones de la máquina tenían una importancia. La circunferencia media unos 630 metros y alcanzaba aproximadamente la mitad de la energía de SLAC a un coste que se acercaba a la décima parte. En Cambridge, Massachusetts, se construyó en 1972 otro sincrotrón de electrones con una energía de haz de 6,3 GeV. Pero se desmanteló. Se instaló un tercero en la ciudad alemana de Hambrugo, el DESY, abreviación de "Deutsches Elektronen-Synchrotron". Con una energía de haz de 7 GeV, DESY se ha convertido en la base para un plurivalente laboratorio de física de altas energías.

Para un radio de curvatura dado, la pérdida de energía por radiación de sincrotrón aumenta con la cuarta potencia de la energía del haz. La pérdida de energía se convierte en el factor dominante en el diseño del acelerador a unos 10 GeV e impone una barrera casi impenetrable después de algunos centenares de GeV. Desde luego la viabilidad de cualquier instrumento es cuestión de decisión y está sujeta a cambio; hace 25 años el límite de sincrotrones de electrones se creía que oscilaba alrededor de 1 GeV.

En defensa de la radiación de sincrotrón, debe indicarse que no es una energía completamente perdida. La radiación produce el amortiguamiento de las pequeñas desviaciones de los electrones con respecto a su trayectoria media, lo que hace que el haz de electrones sea más fácil de controlar que un haz de protones. Y lo que es más importante, la propia radiación se ha convertido en un valioso instrumento para estudios biológicos y de materiales. Constituye la fuente más intensa de radiación ultravioleta y de rayos X, de todas las que se conocen. En varios laboratorios de física de altas energías se han construido servicios para explotar la radiación y se han fabricado, a propósito, varios pequeños aceleradores como fuentes de radiación de sincrotrón.

A pesar del difícil problema de la radiación de sincrotrón, los aceleradores de electrones asumirán una importancia creciente en los futuros programas de física de altas energías. La razón estriba en que el electrón es una partícula mucho más sencilla que el protón y, por tanto, constituye una mejor sonda para estudiar la estructura de la materia. Cuando dos protones (u otros hadrones) chocan, lo que resulta queda complicado por la estructura de muchos quarks de las partículas. El electrón no parece tener, por el momento, una estructura interna, y así los resultados de las interacciones de los electrones se interpretan con más facilidad. Esta pureza se logra en doble medida cuando un electrón choca contra otro electrón o con un positrón. Tales colisiones pueden lograrse construyendo un aparato en el que el haz acelerado incida sobre otro haz de partículas, no sobre un blanco fijo.

Si todos los 500 GeV por partícula que se generan en los mayores sincrotrones de protones se liberaran en una colisión con un blanco fijo, muchas de las metas establecidas para la próxima generación de aceleradores de partículas se habrían logrado ya. Por ejemplo, se espera que el "weakón" tenga una masa próxima a los 100 GeV; se habría visto, por tanto. Sin embargo, no toda la energía de una partícula acelerada está disponible para la creación de partículas nuevas cuando un proyectil en movimiento choca contra un blanco fijo. Por contra, una gran fracción de la energía se destina a poner en movimiento el sistema formado por el par de partículas.

Consideremos un protón de 500 GeV que choque contra un protón en reposo, verbigracia, el núcleo de un átomo de hidrógeno. Si el proyectil se detuviese en virtud de la colisión, toda la energía que había adquirido debería disiparse. Lo que sucede en la realidad es que el protón acelerado sigue moviéndose y el protón del blanco se mueve con él en la misma dirección. A causa de los efectos relativistas, el protón acelerado, que se mueve muy de prisa, tiene una masa 530 veces mayor que el protón del blanco. La partícula acelerada no se para en la colisión, de la misma manera que un camión no se para cuando choca contra una bicicleta. Las dos partículas juntas retienen una energía de 469 GeV y sólo quedan disponibles 31 GeV para la creación de nuevas partículas.

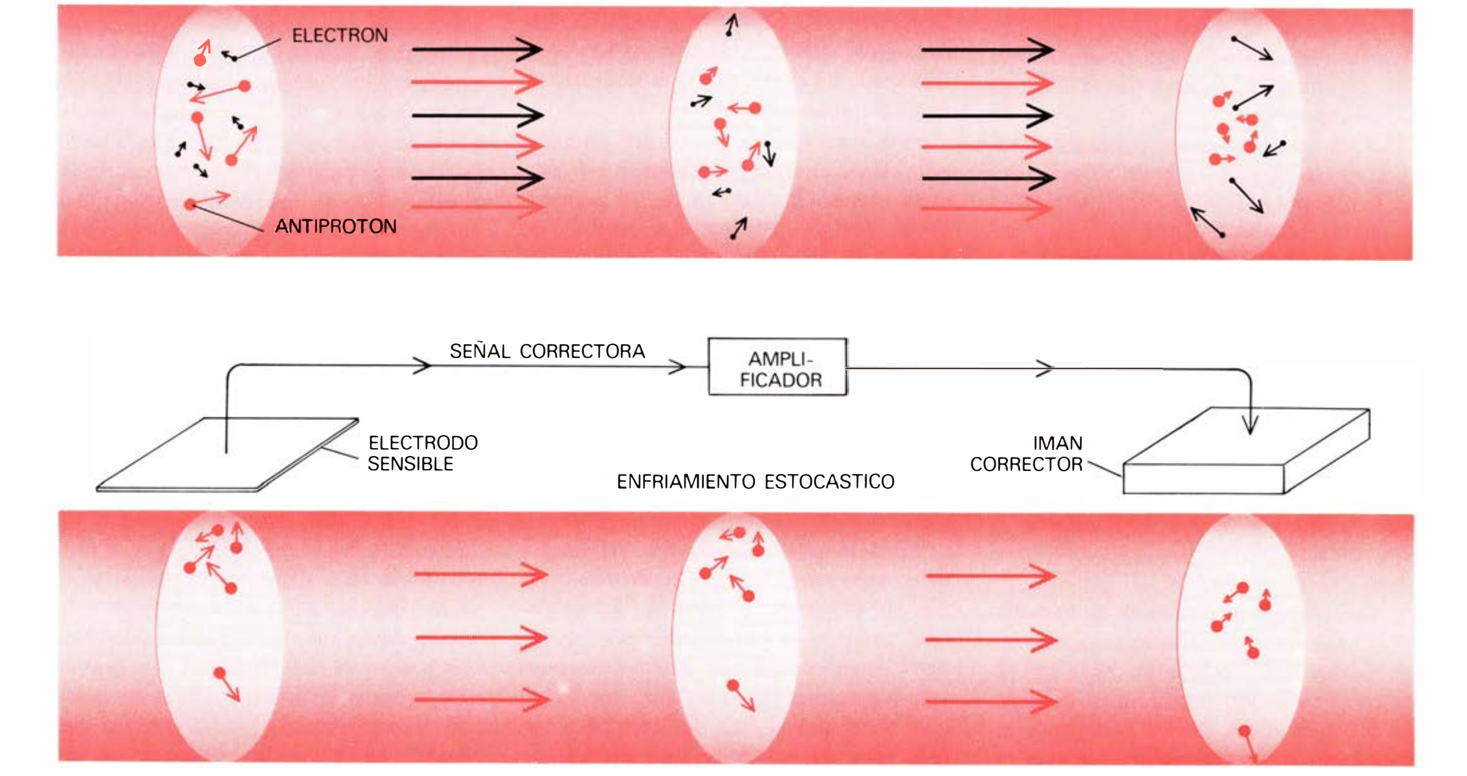
La energía liberada en una colisión

entre partículas puede calcularse más fácilmente cambiando de sistema de referencia. La colisión se apreciaría más realísticamente por un observador que se moviera en dirección paralela al haz y a la misma velocidad que el centro de masas del sistema de dos partículas. Un tal observador vería el protón del haz y el blanco acercándose simétricamente con una energía de 15,5 GeV cada uno, es decir, una energía total de 31 GeV. Esta es la energía del centro de masas. Según predice la cinemática de la teoría de la relatividad, la energía del centro de masas crece sólo en proporción a la raíz cuadrada de la energía del haz, de manera que a medida que los aceleradores aumentan de tamaño se pierde, en este sentido, una parte mayor de la energía invertida. Para alcanzar una energía de centro de masas de 100 GeV con un acelerador de protones de blanco fijo, la energía del haz debería superar los 10.000 GeV.

La visión de la colisión desde el sistema de referencia del centro de masas sugiere una solución a este problema. Si las dos partículas se aceleran a la misma energía, y se mueven en direcciones opuestas, podemos obligarlas a que choquen frontalmente. En la situación más sencilla, ambas podrían quedarse paradas. Así se liberaría toda su energía. Si ambas partículas tienen la misma masa en reposo y si han sido aceleradas a la misma energía, la energía en el sistema de referencia del centro de masas es simplemente la suma de las dos energías de haz. Para alcanzar 100 GeV todo lo que se necesita es un par de haces con una energía de 50 GeV cada uno.

Podemos inducir colisiones frontales construyendo anillos de almacenamiento, en los que que circulan constantemente haces de partículas. El anillo de almacenamiento se parece a un sincrotrón: consta de una cámara de vacío anular que está rodeada por imanes desviadores y focalizadores y tiene, al menos, una cavidad de radiofrecuencia. El anillo no suele emplearse para aumentar la energía de las partículas, sino para mantenerlas en una órbita fija y a una energía constante. La cavidad de radiofrecuencia sólo suministra la potencia necesaria para compensar la pérdida de energía por radiación de sincrotrón.

Un método de lograr colisiones frontales es construir dos anillos de almacenamiento tangentes entre sí. Las partículas chocan en el punto de contacto. En otro esquema, los anillos se entrelazan de manera que cada uno de ellos siga una trayectoria ondulante como los cabos de un bucle trenzado; los haces chocan en cada intersección. El diseño más inteligente de los anillos de almacenamiento sólo es factible cuando los haces que se almacenan están formados por partículas y sus correspondientes antipartículas, tales como electrones y positrones; en cuyo caso llega a bastar un solo anillo. Supongamos que los campos magnéticos y las cavidades de radiofrecuencia se ajusten y mantengan los electrones circulando en el sentido de las agujas del reloj a determinada energía. Los positrones, al tener la misma carga eléctrica pero de signo contrario, responden a los campos eléctricos y magnéticos de manera exactamente opuesta a la de los electrones. Los positrones pueden inyectarse, por tanto, en el mismo anillo, aunque en dirección contraria a la de las agujas del reloj; el sistema único de imanes y de cavidades de radiofrecuencia mantendrá ambas clases de partículas en órbita. Si el anillo contiene un paquete de electrones y otro de positrones, se observarán colisiones en dos puntos diametralmente opuestos.



TECNICAS DE ENFRIAMIENTO DE HAZ, que son esenciales para el éxito de los anillos de almacenamiento protón-antiprotón. Los antiprotones se crean con un rango comparativamente amplio de velocidad y direcciones y, por tanto, no pueden inyectarse directamente en un acelerador. Cuando los antiprotones se miran en su propio sistema de referencia, forman un gas caliente, las partículas del cual tienen velocidades dirigidas al azar. Un medio de enfriar el gas está basado en el empleo de electrones para que se lleven el calor. En repetidas colisiones con los antiprotones, los electrones se van con la mayor parte del momento dirigido al azar. En otro método de enfriamiento (enfriamiento estocástico) se detecta la posición media del haz y entonces se aplica una corrección para mantener el haz centrado en el tubo de vacío. Algunas partículas pueden quedar afectadas adversamente por la corrección, pero la mayoría reducirá sus componentes de movimiento extrañas. Tras muchas repeticiones de este proceso, el haz de antiprotones se enfriará. Se muestra la componente transversal de la velocidad de la partícula; también hay variación de la celeridad a lo largo del eje del haz.

Dada la enorme ventaja energética de un aparato de haces colisionantes, quizá sorprenda que alguien pueda pensar hoy en construir aceleradores de blanco fijo. Sin embargo, la energía no es la única medida que afecta a las características de un acelerador. Otro factor importante es el ritmo al que se observan las interacciones. Un blanco sólido o líquido tiene una densidad de partículas mucho mayor que un haz acelerado; de ello resulta que el ritmo de interacción es mucho mayor que un acelerador de blanco fijo. En el aparato de haces colisionantes, durante la mayor parte del tiempo las partículas de los haces no chocan, sino que pasan unas entre otras sin interaccionar en absoluto. La probabilidad de interacción es proporcional a un parámetro llamado luminosidad, que esencialmente mide el brillo del haz. La luminosidad de las máquinas de blanco fijo es mayor que la de los anillos de almacenamiento en un factor cercano al millón.

La interpretación de los resultados de un experimento físico requiere una muestra de sucesos estadísticamente significativa. Hacerse con esa muestra puede resultar difícil cuando los sucesos de interés acontecen a un ritmo de tan sólo unos pocos por día. Por cuya razón cabe esperar que los anillos de almacenamiento den la primera pista de los fenómenos más corrientes a alta energía; ahora bien, una precisión mayor y la observación de los sucesos raros sólo será posible cuando los aceleradores de blanco fijo alcancen energías de centro de masas comparables.

Otra ventaja de los aceleradores de blanco fijo es su versatilidad. Un anillo de almacenamiento en general proporciona sólo una clase de choques; verbigracia, protones sobre protones o electrones sobre positrones. Una máquina de blanco fijo, por otro lado, puede generar distintos haces secundarios, incluyendo neutrinos, muones, piones y otros mesones, antiprotones y las partículas pesadas llamada hiperones. De hecho, la pequeña fracción de la energía del haz que queda disponible en el sistema centro de masas propicia la creación de haces secundarios; la gran cantidad de energía "perdida" acelera las partículas secundarias.

De otra ventaja de los aceleradores de blanco fijo nos hemos ocupado por extenso: la máquina construida para suministrar partículas con fines experimentales puede suministrarlas más adelante a un acelerador mayor, o a un anillo de almacenamiento.

El primer anillo de almacenamiento electrón-positrón se empezó a construir en 1959 bajo la dirección de Bruno Touschek y un entusiasta grupo de físicos italianos de los Laboratorios Nacionales del Comité Nacional para la Energía Nuclear (CNEN) de Frascati, cerca de Roma. Denominado ADA, abreviación de "Annello d'Accumulazione", este anillo se trasladó más tarde al laboratorio de Orsay en las cercanías de París. La energía de cada haz era de 0,25 GeV, lo que daba una energía en el sistema del centro de masas de 0,5 GeV.

En los primeros años de la década de los 70 operaban ya media docena más de anillos de almacenamiento. El ADA fue el progenitor directo de un anillo mayor de Orsay (luego se construiría allí un tercero) y de un anillo de Frascati que se ha llamado ADONE, derivado de "gran ADA", con una energía de 1,5 GeV por haz. El primero de una sucesión de proyectos pioneros sobre anillos de almacenamiento a realizar en Novosibirsk, Unión Soviética, se terminó en 1965. El Cambridge Electron Accelerator se reformó por las mismas fechas para constituir un anillo de almacenamiento con una energía de 3,5 GeV por haz. La vida de la máquina de Cambridge fue breve, pero reveló una anomalía en el ritmo total de interacciones electrón-positrón que presagió el descubrimiento del encanto.

En 1974, los anillos de almacenamiento de dos laboratorios habían superado ya los 4 GeV por haz. Uno de ellos, llamado SPEAR, está en el SLAC y se llena con electrones y positrones del linac que hay allí. El otro, llamado DORIS, fue construido al lado del sincrotrón DESY, de Hamburgo, y recibe sus partículas de esa fuente.

Al anillo de almacenamiento SPEAR le cabe la gloria de haber posibilitado la observación de la primera partícula encantada, la *J* o psi, a 3,1 GeV. Como ya he dicho, la misma partícula fue descubierta simultáneamente en un experimento de blanco fijo de Brookhaven, pero la señal del SPEAR fue mucho más clara. El personal del DORIS consiguió detectar la nueva partícula inmediatamente cuando sintonizaron su máquina a la energía correcta. El ADONE perdió la sin par ocasión de ver la psi debido a que se había diseñado para una energía máxima de centro de masas de 3 GeV, pero fue posible forzar la máquina por encima de este límite y detectar la partícula. La subsiguiente exploración del espectro de las partículas encantadas se realizó principalmente con estos tres anillos de almacenamiento.

El descubrimiento del encanto revitalizó la física de las partículas elementales. Obviamente, satisfizo a los experimentadores que habían hecho el descubrimiento, pero también fue recompensador para los físicos teóricos, quienes habían predicho la existencia del quark encantado y habían esperado diez años a que se manifestara. Tras un éxito de ese calibre no maravilla demasiado que se hicieran inmediatamente planes para una nueva ronda de anillos de almacenamiento electrón-positrón. Pronto fue aprobada una propuesta para construir un anillo mayor en el SLAC, en colaboración con el Lawrence Berkeley Laboratory; heredó el nombre de PEP. En Europa se consideraron dos proyectos similares, uno en Hamburgo, que se llamaría PETRA, y el otro en el Rutherford Laboratory, Inglaterra. Las dos propuestas fueron presentadas ante un comité internacional, el Comité Europeo para Futuros Aceleradores; siguiendo la recomendación del comité, se empezó la construcción del proyecto alemán y se desestimó el plan del Rutherford.

Las características del PETRA y del PEP muestran un extraordinario parecido. El PETRA está diseñado para alcanzar una energía máxima de 19 GeV por haz, el PEP, de 18 GeV. En tamaño (2,2 kilómetros de circunferencia) y luminosidad son casi idénticos. Lo más probable es que los costes sean similares, aunque las diferencias en los métodos de estimación hacen que la comparación resulte difícil.

En cuanto se inició la construcción quedó de manifiesto que los dos laboratorios estaban en una carrera. PETRA ha ganado. Empezó a funcionar en 1978, mientras que no se espera que el PEP esté en actividad hasta inicios de este año. El avance más rápido del PETRA puede atribuirse al elegante y sobrio diseño de Gustav Adolph Voss, al hecho de que el grupo alemán empezó antes y quizá también a que dispusieron de dinero a medida que lo necesitaron. Se animó a los constructores del PETRA a gastar rápidamente a fin de estimular los negocios en el área de Hamburgo. Por contra, el PEP se ha alimentado de fondos suministrados dólar a dólar y también se ha retrasado por fuertes lluvias (que inundaron las excavaciones del túnel) y por problemas laborales.

Los primeros experimentos del PETRA permiten abrigar la esperanza de que a ambas máquinas les aguarda un futuro prometedor. Los primeros resultados proporcionan una base alentadora para la cromodinámica cuántica, la teoría nacida para describir la interacción de los quarks y gluones dentro de los hadrones.

En la carrera por explotar la sencillez de las colisiones electrón-positrón a altas energías no participan sólo PETRA y PEP. Acaba de entrar en el hipódromo

un caballo misterioso. Es el Cornell Electron Storage Ring, o CESR, encajado en el mismo túnel que el sincrotrón de electrones existente en Cornell. Bajo la dirección de Boyce D. McDaniel, de Cornell, el proyecto se ha acabado en menos de dos años y a un coste inferior a los 20 millones de dólares. Las primeras colisiones electrón-positrón se detectaron el último verano y ya se han iniciado los primeros experimentos. El CESR se sale de lo corriente al estar financiado por la National Science Foundation; prácticamente todos los otros laboratorios de física de altas energías de los Estados Unidos reciben la subvención económica de las arcas del Departamento de Energía.

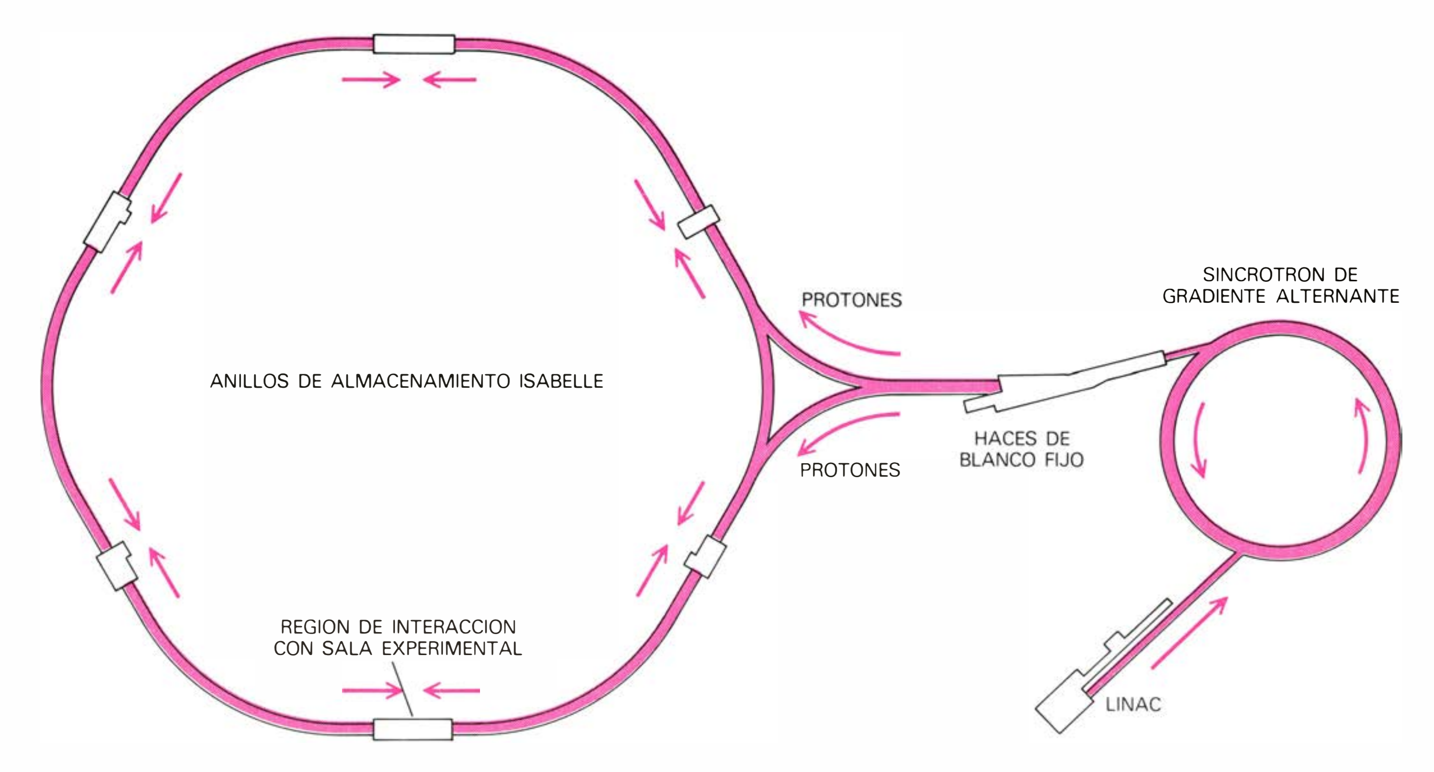
Maury Tigner, de Cornell, introdujo en el CESR un nuevo método de llenar el anillo de almacenamiento. Se pasan repetidamente paquetes de partículas entre el CESR y el sincrotrón de electrones, que tienen radios distintos. Aprovechando el ligero desfase del anillo exterior pueden superponerse muchos paquetes en el anillo de almacenamiento, mejorando por tanto la luminosidad de los haces. La energía máxima es de 8 GeV por haz, que puede convertirse en una elección afortunada. Intermedia entre las energías de SPEAR y DORIS y las de PEP y PETRA, cae en una zona en la que se espera encontrar un rico espectro de nuevas partículas. Estas partículas son las relacionadas con la hypsilón, que incorporan el quark fondo ("bottom").

Animados por el éxito, el grupo de Cornell ha hecho el diseño inicial de un anillo de almacenamiento electrón-positrón de 50 GeV por haz. Con una energía de centro de masas de 100 GeV, una perspectiva sugestivamente excitadora será la producción abundante del weakón neutro, o Z^0. A esa energía, los electrones y positrones radiarían 1,6 GeV de radiación de sincrotrón en cada vuelta, pérdida de energía que sería compensada por cavidades de radiofrecuencia superconductoras con un consumo energético total de 36 megawatt. La máquina, de cuatro kilómetros de circunferencia, costaría 120 millones de dólares. En fase de proyecto adelantado se encuentra otro aparato electrón-positrón, todavía mayor. Lleva el nombre de LEP, abreviación de Gran Anillo de Almacenamiento Electrón-Positrón ("Large Electron-Positron"); fue propuesto por físicos del CERN. En 1978 se presentaron, ante el Comité Europeo para Futuros Aceleradores, varios candidatos para hacerse con la sede de la magna instalación europea destinada a la física de altas energías. La autorización del Comité la ganó el LEP. En seguida se emprendió un estudio detallado del proyecto. No se han asignado fondos para la construcción del anillo, pero en cuanto exista presupuesto para un gran proyecto europeo, cabe esperar que el LEP tenga todas las prioridades.

En una primera fase, la energía máxima del LEP podría ser de unos 86 GeV por haz, pero podría aumentarse después hasta 130 GeV añadiendo cavidades de radiofrecuencia superconductoras. A fin de minimizar la pérdida de energía por radiación de sincrotrón se construiría un anillo muy grande, con una circunferencia de 30 kilómetros. Como en el caso del SPS, se excavaría un túnel subterráneo. Habría de atravesar, por debajo, las montañas del Jura al oeste de Ginebra y tres de las zonas experimentales estarían a unos 800 metros de profundidad. Colocando el LEP junto al SPS podrían añadirse, más adelante, dispositivos para choques entre electrones de un anillo y protones del otro. En realidad, una vez que esté abierto el túnel del LEP se podría agregar un anillo de imanes superconductores para un sincrotrón de protones. Los protones inyectados por el SPS podrían llevarse a una energía entre 3 y 6 TeV, aunque de semejante posibilidad nada han dicho en público los proponentes de la máquina.

El coste estimado de la fase inicial del LEP es de 1275 millones de francos suizos, más que el coste conjunto del Fermilab y del SPS del CERN. Si la construcción empezara pronto, el anillo podría entrar en funcionamiento en las postrimerías de nuestra década.

No hay razón para que los aparatos de haces colisionantes se limiten a electrones y positrones. Desde 1971 viene funcionando en el CERN una instala-



LAS COLISIONES PROTON-PROTON son la finalidad del ISABELLE, un aparato que se está construyendo en Brookhaven. Los protones serán suministrados por el Sincrotrón de Gradiente Alternante, un acelerador de 33 GeV construido en 1961. Se inyectarán en anillos entrelazados formados por imanes superconductores y llevados a 400 GeV por haz; por tanto, la energía en el sistema del centro de masas será de 800 GeV. Se tiene especial cuidado en asegurar una luminosidad elevada, 10 veces mayor que la de cualquier anillo de almacenamiento. El ISABELLE se terminará en 1986.

ción para producir colisiones protón-protón. Recibe el nombre de Anillos de Almacenamiento Interseccionantes, "Intersecting Storage Rings" (ISR); consta de dos anillos entrelazados que almacenan haces de protones que giran en sentidos opuestos. Los anillos se cortan en ocho puntos alrededor de su circunferencia formando en todos ellos un ángulo pequeño. Se pueden colocar detectores en siete de las zonas de intersección.

Los protones almacenados en el ISR los suministra el PS adyacente, a energías de hasta 28 GeV. En el ISR pueden acelerarse ligeramente hasta una energía de 31 GeV por haz. La energía de centro de masas resultante, 62 GeV, es la máxima disponible actualmente. Para alcanzar la misma energía de centro de masas con un acelerador de blanco fijo se necesitaría una energía de haz de 2 TeV.

Un instrumento de concepción similar al ISR, aunque mucho mayor en tamaño, se halla en fase de construcción en Brookhaven. Se llama ISABELLE (por "Intersecting Storage Accelerator", Acelerador de Almacenamiento Interseccionante, y "belle", o bello). Constará de dos anillos entrelazados, que serán unas cuatro veces mayores que los del ISR. Además, ambos anillos estarán formados por imanes superconductores con una intensidad de campo mucho mayor que la de los imanes convencionales. Los anillos funcionarán a 400 GeV por haz, con una energía en el sistema del centro de masas de 800 GeV. El acelerador equivalente de blanco fijo necesitaría una energía de haz de 340 TeV.

Los protones para el ISABELLE vendrán del AGS. Se inyectarán a unos 30 GeV; cada uno de los anillos del ISABELLE funcionará entonces brevemente, como un sincrotrón para llevar los protones a la energía máxima. Las corrientes de los imanes se mantendrán fijas y los protones circularán con energía constante. Hay seis intersecciones y dispone de espacio adecuado para la experimentación en todas ellas.

Los imanes superconductores difieren en varios detalles de los que se construyen en el Fermilab para el tevatrón. Por ejemplo, la cámara de vacío que pasa a través de los imanes superconductores del tevatrón se enfriará a la temperatura del helio líquido a la que están los imanes, mientras que el tubo de vacío del ISABELLE estará tibio. Por otro lado el brazo de hierro comparativamente pesado de los imanes del ISABELLE se enfriará mientras que el brazo del tevatrón permanecerá a temperatura ambiente. Los propios sistemas de enfriamiento serán muy diferentes. El tevatrón empleará helio líquido, mientras que el ISABELLE se enfriará por la circulación forzada de helio de alta densidad a una presión por encima del punto crítico, donde la distinción entre líquido y vapor desaparece. Por haber tanto hierro en los imanes del ISABELLE, se presume que se tardará dos semanas en enfriarlos a su temperatura de trabajo de 3,8 grados Kelvin.

ACELERADOR	FECHA DE PUESTA EN MARCHA	PARTICULAS ACELERADAS	ENERGIA DEL HAZ (GeV)	ENERGIA CENTRO DE MASAS (GeV)	INTENSIDAD (PARTICULAS POR PAQUETE)
PS CERN, Ginebra	1959	Protones	28	7.4	
AGS Brookhaven National Laboratory, New York	1961	Protones	33	8	8×10^{12}
SLAC (acelerador lineal) Stanford University	1961	Electrones	22	6.5	
Cornell Electron Synchrotron Cornell University	1967	Electrones	12	4.9	
Serpukhov Proton Synchrotron Serpukhov, Unión Soviética	1967	Protones	76	12	5×10^{12}
Fermilab, anillo principal Batavia III	1972	Protones	500	30.7	2×10^{13}
DESY Hamburgo	1974	Electrones	7	3.8	
KEK Japón	1975	Protones	12	5	
SPS CERN, Ginebra	1976	Protones	500	30.7	10^{13}
Tevatron Fermilab	1982	Protones	1,000	43	5×10^{13}
Beijing Proton Synchrotron Beijing	1985	Protones	50	9.8	10^{13}
UNK Serpukhov, Unión Soviética	Finales década 1980	Protones	3,000	75	
Pentevac Fermilab	Finales década 1980	Protones	5,000	97	10^{14}
VBA Lugar no seleccionado	No programada	Protones	20,000	137 to 194	10^{13}–10^{15}

INVENTARIO DE LOS GRANDES ACELERADORES, que incluye un linac (en el SLAC); el resto son sincrotrones. Entre los proyectos no acabados están en fase de construcción el tevatrón del Fermilab y el Sincrotrón de Protones de Beijing en China. El UNK, un sincrotrón de protones de 3 TeV que se planea construir en Serpukhov, cerca de Moscú, no tiene aún la autorización final. El Pentevac es puro diseño, que describe el acelerador más potente que podría construirse en la zona del Fermilab. Se ha propuesto el VBA (Muy Gran Acelerador o hiperacelerador) como sucesor de los aceleradores y anillos de almacenamiento planeados o en construcción hasta la fecha. Su máxima energía y su coste serían un orden de magnitud mayor que los de los mayores aceleradores que funcionan actualmente. La construcción del VBA requerirá los recursos conjuntos de Estados Unidos, la URSS y otras naciones.

La abertura de los imanes del ISABELLE, será mayor. Permitirá que los atraviese una cámara de vacío de más calibre; lo cual, a su vez, posibilitará haces de protones más intensos y mayor luminosidad. La meta es una luminosidad diez veces mayor que la de cualquier aparato de haces colisionantes existente y 25 veces mayor que la luminosidad del ISR. El valor de ISABELLE se ha estimado en 275 millones de dólares. Los planes de construcción prevén que los experimentos empiecen en 1986.

Antes de que se acabe el ISABELLE se habrán logrado, quizás, energías de centro de masas todavía mayores en el Fermilab y en el CERN, aunque las luminosidades serán mucho más bajas. Las máximas energías se habrán obtenido mediante modos improvisados de funcionamiento de los grandes sincrotrones como anillos de almacenamiento. En el Fermilab la presencia de dos anillos en el mismo túnel hace posibles varias disposiciones de haces colisionantes. Un plan sería almacenar haces de protones girando en sentido contrario en los dos anillos y hacerlos chocar en una de las partes rectas que se reparten alrededor de los anillos. El tevatrón podría

INSTRUMENTO	FECHA DE PUESTA EN MARCHA	TIPO	PARTICULAS ALMACENADAS	ENERGIA DE CADA HAZ (GeV)	ENERGIA DE CENTRO DE MASAS (GeV)	LUMINOSIDAD (CM^{-2} SEG^{-1})
ADA (desmantelado) Laboratorios nacionales del CNEN, Frascati	1963	Anillo sencillo	e^+e^-	.25	.5	
Princeton Stanford Storage Rings (desmantelado) Stanford University	1964	Anillos tangentes	e^-e^-	.56	1.1	10^{27}
VEPP-2 Novosibirsk, Unión Soviética	1965	Anillos tangentes	e^+e^-	.04	.08	
ACO Laboratorios de Orsay, París	1966	Anillo sencillo	e^+e^-	.5	1	
ADONE Laboratorios nacionales del CNEN, Frascati	1965	Anillo sencillo	e^+e^-	1.5	3	
ISR CERN, Ginebra	1971	Anillos entrelazados	pp	31	62	4×10^{31}
CEA-Bypass (desmantelado) Cambridge, Mass	1967	Anillo sencillo	e^+e^-	3.5	7	2×10^{28}
SPEAR Stanford Linear Accelerator Center	1972	Anillo sencillo	e^+e^-	4.2	8.4	10^{32}
DORIS DESY, Hamburgo	1974	Anillo sencillo	e^+e^-	4.5	9	10^{32}
VEPP-2M Novosibirsk, Unión Soviética	1975	Anillo sencillo	e^+e^-	1.3	2.6	
DCI Laboratorios de Orsay, París	1975	Anillos entrelazados	e^+e^-	3.7	7.4	10^{32}
VEPP-3 Novosibirsk, Unión Soviética	1977	Anillo sencillo	e^+e^-	3	6	10^{30}
VEPP-4 Novosibirsk, Unión Soviética	1978	Anillo sencillo	e^+e^-	7	14	10^{31}
PETRA DESY, Hamburgo	1978	Anillo sencillo	e^+e^-	19	38	10^{32}
CESR Cornell University	1979	Anillo sencillo	e^+e^-	8	16	10^{32}
ISR $p\bar{p}$ CERN, Ginebra	1980	Anillos entrelazados	$p\bar{p}$	31	62	10^{29}
PEP Stanford Linear Accelerator Center	1980	Anillo sencillo	e^+e^-	18	36	10^{32}
SPS $p\bar{p}$ CERN, Ginebra	1981	Anillo sencillo	$p\bar{p}$	270	540	10^{30}
Fermilab $p\bar{p}$ Batavia, Ill	1982	Anillo sencillo	$p\bar{p}$	1,000	2,000	10^{30}
VAPP Novosibirsk, Unión Soviética	?	Anillo sencillo	$p\bar{p}$	23	46	
ISABELLE Laboratorio nacional de Brookhaven	1986	Anillos entrelazados	pp	400	800	10^{33}
LEP CERN, Ginebra	Finales década 1980	Anillo sencillo	e^+e^-	86	172	10^{32}
UNK Serpukhov, Unión Soviética	Finales década 1980	Anillo sencillo	$p\bar{p}$	3,000	6,000	
VBA Lugar no seleccionado	No programada	Anillo sencillo	$p\bar{p}$	20,000	40,000	10^{32}

ANILLOS DE ALMACENAMIENTO de partículas. Tendrán un papel importante en el programa de la física de alta energía de las dos próximas décadas. Dos nuevos anillos electrón-positrón han empezado a funcionar recientemente; una tercera de tales máquinas se acabará pronto. Se trata del CESR, un anillo de 8 GeV de la Universidad de Cornell, PETRA y PEP; los dos últimos aparatos presentan características casi idénticas, y pertenecen al acelerador de Hamburgo y del SLAC, respectivamente. Cerca de una docena más de aparatos de haces colisionantes están planeados, incluyendo numerosas propuestas de hacer funcionar sincrotrones como anillos de almacenamiento. El mayor proyecto para el cual existe un diseño detallado es el LEP, el Gran anillo de almacenamiento Electrón-Positrón ("Large Electron-Positron"), que se construiría en el CERN (Ginebra, Suiza). La fase inicial del LEP tendría una energía de 86 GeV por haz, pero podría elevarse a 130 GeV con cavidades de radiofrecuencia superconductoras.

funcionar a su máxima energía de 1 TeV como un anillo de almacenamiento, pero la máxima corriente estacionaria de los imanes limitaría el anillo principal a unos 250 GeV. La energía de centro de masas para tales colisiones se acerca a 1 TeV. Haciendo pulsar el anillo principal hasta su máxima energía se podría obtener una mejora de la energía en el sistema del centro de masas de un 40 por ciento, pero ello resultaría en notable detrimento de ritmo de las interacciones.

Una propuesta que daría aún mayor energía y que por fin se considera seriamente es el empleo del tevatrón como un aparato de haces colisionantes de un solo anillo para protones y antiprotones, funcionando según los mismos principios que los anillos electrón-positrón. La dificultad de este proyecto reside en la acumulación de un haz de antiprotones suficientemente intenso, ya que éstos no existen en la materia ordinaria. Los antiprotones se forman en las colisiones de protones con un blanco fijo, pero se producen con un amplio espectro de velocidades y direcciones y no pueden inyectarse en tal estado dentro de la estrecha abertura de un sincrotrón. Contemplados desde el sistema de referencia de los antiprotones, el haz resultante de las colisiones es un gas caliente, cuya temperatura debe reducirse sustancialmente antes de que las partículas puedan emplearse en un experimento de haces colisionantes.

Los planes para construir un anillo de almacenamiento protón-antiprotón dependen de dos métodos, de reciente desarrollo, para enfriar artificialmente un haz de antiprotones. El primer método, conocido por enfriamiento por electrones ("electron cooling"), fue inventado por el fallecido Gersh I. Budker y sus colegas de Novosibirsk. Se confina un haz cualquiera de antiprotones en un anillo de almacenamiento de baja energía y muy gran abertura. Se hace pasar luego electrones por un largo tramo longitudinal del anillo, de suerte que su movimiento sea paralelo a la trayectoria media de los antiprotones y tengan la misma velocidad media. Los electrones ofrecen una temperatura mucho más baja; debido a las colisiones se llevan las componentes, dirigidas al azar, del momento de los antiprotones. A raíz de lo cual, el gas caliente de antiprotones cede su calor al gas frío de electrones. Al final del tramo recto las dos clases de partículas se separan mediante un campo magnético. Los electrones inciden sobre un colector y los antiprotones continúan circulando y pasan, repetidamente, a través de la región de enfriamiento.

El enfriamiento a través de electrones fue demostrado por Budker. Se han construido anillos de enfriamiento mayores en el CERN y en el Fermilab. El personal del CERN fue el primero en repetir los resultados de Novosibirsk; observaron que la velocidad de enfriamiento superaba incluso a la esperada. Los planes del Fermilab contemplan la inyección de protones y antiprotones enfriados en el tevatrón a una energía de unos 150 GeV. Si se quiere inyectar ambos haces será preciso disponer de un método bastante complicado de transmitir las partículas entre el anillo principal, el tevatrón, el acelerador intermedio y el anillo de enfriamiento. Las partículas y antipartículas se acelerarán entonces simultáneamente hasta una energía máxima de 1 TeV. Así la energía en el sistema del centro de masas puede alcanzar dos TeV.

En el CERN se está probando otro método de enfriamiento de antiprotones. Inventado por Simon Van der Meer, se reconoce por enfriamiento estocástico. Los antiprotones también se almacenan en un pequeño anillo de gran abertura, en este caso equipado con un sistema de sensores electrónicos y un imán especial de corrección de órbita. En cada vuelta, los sensores detectan la posición media de las partículas en algún punto a lo largo del anillo. Si el centro de carga de una fluctuación de las partículas del haz se ha separado del eje de la cámara de vacío, se calcula una señal correctora y se envía al imán corrector. (Debido a que los antiprotones se mueven a una velocidad cercana a la de la luz, la señal puede llegar al imán antes que las partículas con sólo seguir una trayectoria más corta a lo largo de una cuerda de la circunferencia.) El imán corrector restablece el centro de carga a su posición correcta. En el proceso cabe que algunas partículas se desvíen aún más de su trayectoria correcta en vez de volver a ella, pero la mayoría siempre se mueve en la buena dirección. El sistema de sensores e imanes actúa como un demonio de Maxwell enfriando el haz mediante hábiles manipulaciones de la distribución de las velocidades de las partículas.

En el CERN los protones y antiprotones se almacenarán en el SPS. Debido a las limitaciones de la corriente estacionaria de los imanes convencionales, la máxima energía mantenida será de 270 GeV por haz. Esto representa 540 GeV en el sistema del centro de masas. Haciendo pulsar los imanes, podrá alcanzarse una energía en el sistema del centro de masas de 1000 GeV, aunque a un ritmo medio de interacciones mucho menor.

Los esfuerzos paralelos del CERN y del Fermilab para establecer colisiones protón-antiprotón se contemplarán como otra carrera de competición, y será con cierta razón. El premio en disputa bien pudiera ser la primera observación del "weakón". El CERN ha dedicado gran parte de sus recursos al proyecto; aunque de momento lleva una ventaja considerable, las fechas estimadas para la finalización, a principios de 1981 para el CERN y uno o dos años más tarde para el Fermilab, están lo suficientemente cercanas para que factores imprevistos puedan alterar el orden de la carrera. En cualquier caso, hemos de confesar que existe un espíritu de generosa cooperación entre los dos laboratorios. Ambos proyectos han sido muy inspirados por las ideas brillantes del mismo hombre: Carlo Rubbia, de la Universidad de Harvard.

No se espera que ninguno de estos anillos de almacenamiento improvisados tenga una gran luminosidad. Las posibilidades de aparatos experimentales serán también menores que lo ideal. Además habrá una reñida competencia en lo que respecta al tiempo de uso con programas, muy atractivos, de blanco fijo, que son incompatibles con los experimentos de haces colisionantes. La atracción irresistible de los proyectos de haz enfriado reside en la oportunidad de alcanzar energías extraordinariamente altas de manera rápida y barata. Los colisionadores ("colliders") del CERN y del Fermilab servirán para mirar, desde el Monte Pisga, la tierra prometida. ¿Nos llevarán allí el ISABELLE y el LEP?

No debe quedar la impresión de que sólo haya laboratorios interesantes de física de altas energías en los Estados Unidos y en Europa Occidental. En el Japón ha estado funcionando varios años un sincrotrón de protones de 12 GeV, el KEK, y está planeado un sincrotrón de 200 GeV, el TRISTAN. Los chinos han anunciado planes de un ambicioso complejo de aceleradores a construir en un lugar a 60 kilómetros al noroeste de Beijing. La construcción ha empezado con el inicio de un sincrotrón de protones de 50 GeV que esperan terminar hacia 1985. El maravilloso lugar, desde el que se contemplan las tumbas Ming y que está cerca de la Gran Muralla, es lo suficientemente grande para acomodar un sincrotrón de protones superconductor de 1,5 TeV para el cual la máquina de 50 GeV podría servir como inyector.

El panorama de los laboratorios de aceleración en la Unión Soviética es comparable al de los programas de América y Europa. Una serie de cuatro

aparatos de haces colisionantes electrón-positrón se ha construido en Novosibirsk, el último de los cuales tiene una energía de haz de 7 GeV. Existe un sincrotrón de electrones de 6 GeV en Yerevan, Armenia, y hay pequeños sincrotrones de protones, en Moscú (7 GeV), y en el Instituto Conjunto para la Investigación Nuclear, en Dubna (10 GeV). El Instituto de Serpukhov para Física Nuclear, a cien kilómetros al sur de Moscú, tiene un acelerador que durante varios años fue el más potente del mundo: se trata de un sincrotrón de protones de 76 GeV. Cuando se finalice una ampliación de los dispositivos que se ha propuesto el laboratorio de Serpukhov podrá reclamar de nuevo aquella distinción.

En el nuevo proyecto, que se llama UNK, se transferirán los protones del acelerador de 76 GeV a un nuevo lugar a seis kilómetros de distancia. Allí los protones se inyectarán en un sincrotrón de 400 GeV, que servirá para empujarlos a un tercer sincrotrón, a construir en el mismo túnel, con una energía máxima de 3 TeV. El sincrotrón intermedio será de diseño convencional, pero el anillo de 3 TeV usará imanes superconductores. Los imanes se hallan en fase de desarrollo en Serpukhov y en el Centro de Investigación Nuclear de Saclay.

Se contemplan varios modos de funcionamiento. Los protones de 3 TeV y los haces disponibles para experimentos estarán disponibles para experimentos de blanco fijo. También serán posibles las colisiones de los protones de 3 TeV y los protones de 400 GeV del segundo acelerador, lo que dará una energía de centro de masas algo mayor que 1 TeV. Se ha examinado asimismo la posibilidad de añadir un sincrotrón de electrones de 20 GeV que proporcionaría colisiones electrón-protón. El enfriamiento del haz proporcionaría antiprotones para colisiones con una energía en el sistema del centro de masas de 6 TeV. Si se construyera un segundo gran anillo se podría lograr la misma energía en colisiones protón-protón. Los planes y características del UNK se están perfilando; se tardará unos siete años en finalizar la máquina.

Un sueño que ni siquiera ha llegado a abrirse camino a través de la burocracia es el acelerador "Site-Filler" o Pentevac, que podría emplazarse en el Fermilab. La idea surgió a raíz de la pregunta: ¿Cuál es el acelerador más potente que podría instalarse en la zona del Fermilab? El terreno disponible permitiría un acelerador no mayor de cinco kilómetros de diámetro, o 17 kilómetros de circunferencia. La energía que podría obtenerse en un sincrotrón de protones de tal tamaño depende en gran medida de los futuros desarrollos en la tecnología de los imanes superconductores. La máxima intensidad de campo posible en este momento en un imán adecuado para un acelerador es de unos 50 kilogauss. Parece razonable suponer que en unos 10 años, con nuevos materiales y nuevos métodos de fabricación, el límite aumentará a unos 85 kilogauss (aunque quizá se obtengan campos mucho mayores). En dicho caso el acelerador podría llegar a 5 TeV. Un variado menú de experimentos sería posible entonces. En la lista se incluirían haces de blanco fijo, colisiones de protón-protón a energías de hasta 4,5 TeV y colisiones protón-antiprotón a 10 TeV. Podría añadirse un acelerador de electrones, aumentando aún más las posibles combinaciones del laboratorio.

Quizá alguien defienda la tesis de que el mundo necesita dos sincrotrones muy parecidos, como el anillo principal del Fermilab y el SPS, o dos anillos de almacenamiento muy similares, como el PEP y el PETRA. Lo cierto es que ninguno de esos instrumentos permanecerá parado por falta de interesantes experimentos a realizar. Para los mayores aceleradores ahora considerados, sin embargo, cualquier posible ventaja en su duplicación puede someterse a examen. El presupuesto mundial para la física de altas energías quizá no permita más de un sincrotrón de protones con una energía de haz de 20 o 30 TeV. Si alguna vez se impone construir dicha máquina, no podrá hacerse como un proyecto nacional, ni es probable que se emprenda por un consorcio regional como el CERN. La cooperación mundial sería la adecuada para semejante empeño; quizá llegara a ser necesaria. Los físicos sueñan en que sea posible.

Con una curiosa mezcla de pragmatismo y utopía los investigadores han predicado y practicado, desde hace tiempo, la colaboración internacional. El CERN es un modelo ejemplar, pero hay otros esfuerzos internacionales considerables, algunos con una composición más amplia incluso. En 1961, como resultado del movimiento Atomos para la Paz, se hizo un intercambio de físicos de partículas entre los Estados Unidos y la Unión Soviética para el análisis de asuntos de interés mutuo. Entre éstos, la posibilidad de un "Acelerador Mundial" financiado por los recursos unidos de muchas naciones. Las conversaciones acabaron con el episodio del vuelo del *U*-2 sobre la Unión Soviética, pero la idea no murió.

El entusiasmo por el Acelerador Mundial surgió de nuevo en 1975 en una reunión internacional convocada en Nueva Orleans por Victor F. Weisskopf, del Instituto de Tecnología de Massachusetts. Lo que fue un sentir general en la reunión de Nueva Orleans se ha convertido hoy en un ente oficial, el Comité Internacional para Futuros Aceleradores, que opera bajo los auspicios de la Unión Internacional de Física Pura y Aplicada. El comité está encargado de coordinar los planes de los distintos laboratorios nacionales e internacionales, así como de promover la colaboración entre ellos. Muchos físicos piensan que tales proyectos conjuntos ya funcionan bien, y recelan de la mediación oficial.

Mayor entusiasmo ha recibido una segunda función del Comité. Se trata de propiciar un fórum para la discusión de lo que se llama actualmente el VBA ("Very Big Accelerator"). Con tal fin ya se han celebrado dos reuniones de trabajo internacionales, una en el Fermilab en 1978 y la segunda en el CERN hace algunos meses. Ambas reuniones afrontaron las posibilidades y los límites de la tecnología de los aceleradores y los detectores de partículas, aunque, obviamente, las conversaciones tuvieron lugar en el contexto de los fenómenos físicos que son de esperar a las energías del VBA.

Según parece, la construcción de un sincrotrón de protones de 20 TeV es factible. Y no requerirá que se varíe significativamente la práctica actual. Una gran parte del aumento de energía se alcanzaría con la estrategia más sencilla: aumentando el radio del anillo. Por otro lado, los aparatos de haces colisionantes de muy alta energía se podrían beneficiar de nuevas técnicas. Para colisiones electrón-positrón, por ejemplo, se podrían alcanzar energías superiores a las del rango del LEP con una máxima eficiencia mediante un par de linacs enfrentados o mediante una combinación de linacs y sincrotrones de electrones. Se ha sugerido ya un medio de aumentar la luminosidad de los haces colisionantes sin acelerar más partículas. Consiste en aumentar la densidad de partículas de los haces reduciéndolos a una sección del haz a tamaño microscópico en el lugar de la interacción. Esto presenta un nuevo fenómeno que apareció en la primera reunión de trabajo mantenida por el Comité Internacional para Futuros Aceleradores: en un haz pequeño, el campo magnético crece, de manera que la radiación de sincrotrón en el propio campo magnético se convierte en una pérdida importante de energía. Este fenómeno se conoce por "beamstrahlung"

o radiación de haz, variación humorística de "bremsstrahlung", o radiación de frenado, proceso por el cual las partículas cargadas emiten radiación electromagnética al frenarse. Burton Richter y sus colaboradores de SLAC proponen con ahínco que la colisión de haces de linacs de 50 GeV podría constituir una fábrica de Z^0, un desafío a los planes de Cornell.

Una cosa en la que están de acuerdo los físicos involucrados en este campo es que no debe acudirse a un consorcio mundial hasta que no se reconozca claramente que es el único modo de lograr la construcción de un determinado acelerador. Resulta significativo que cuando se planteó con seriedad, en 1960, la posibilidad de un Acelerador Mundial, se tratase de un acelerador de protones de 100 GeV. Ahora, debido a los avances tecnológicos, pueden alcanzarse energías 10 veces mayores con sólo un esfuerzo nacional. Cuando se organizó el Comité Internacional para Futuros Aceleradores parecían proyectos apropiados para un esfuerzo mundial un sincrotrón de protones de 10 TeV y un anillo de almacenamiento electrón-positrón con haces colisionantes de 100 GeV. La propuesta de construcción del LEP en el CERN se ha anticipado a la última posibilidad. Las reuniones de trabajo organizadas por el comité han llegado a la conclusión de que sería posible construir anillos de almacenamiento de electrones con una circunferencia de unos 100 kilómetros y una energía de unas dos veces la del LEP. Por encima de esta energía, los haces colisionantes de los linacs con energías de hasta 350 GeV parecen ser más factibles. Para protones, un sincrotrón con una circunferencia entre 50 y 100 kilómetros, con energía en la zona de los TeV y con todas las opciones de colisiones protón-antiprotón constituiría un firme candidato para el VBA (hiperacelerador).

Si hubiera que recurrir a la construcción de un acelerador tan grande para estudiar la estructura interna de leptones y quarks, y probablemente será necesario hacerlo, competirá al Comité Internacional para Futuros Aceleradores aportar la base internacional precisa que permita la puesta en obra de dicho esfuerzo. Me viene a la memoria una fantástica sugerencia que hiciera 30 años atrás Enrico Fermi: construir un acelerador que diera la vuelta al mundo. De acuerdo con ese patrón, el acelerador tipo VBA sería insignificante, pero el esfuerzo común necesario para construir una máquina así puede ayudar a unir todas las voluntades del mundo.

Biología celular del envejecimiento humano

Al cultivar in vitro células humanas normales, éstas acaban por perder su capacidad funcional y dejan de dividirse. El tiempo en que esto sucede podría representar el límite máximo de duración de la vida humana

Leonard Hayflick

La duración de la vida humana no ha aumentado de forma significativa durante milenios. Para la mayoría de las personas los noventa años bíblicos constituyen aún el límite superior. Lo que ha cambiado, al menos en los países desarrollados, es la esperanza de vida global. La mejora de las condiciones de vida y los avances de la medicina, en particular el control de las enfermedades infecciosas, tan sólo han permitido que un número mayor de personas alcance lo que parece ser el límite máximo de duración de vida. Incluso si se eliminaran las principales causas de muerte, la duración de la vida humana se mantendría en unos 90 o 100 años. La explicación radica en que, aun cuando la muerte pueda sobrevenir como consecuencia de enfermedades cardiovasculares o del cáncer, tales enfermedades derivan de la normal degeneración por envejecimiento de las paredes arteriales o del sistema inmunitario. La causa fundamental de muerte en los últimos años de la vida es el incremento de la vulnerabilidad del organismo frente a enfermedades o accidentes, vulnerabilidad atribuible al declinar inexorable de la capacidad funcional a partir de los 30 años. Bernard L. Strehler, de la Universidad de California del Sur, ha estimado que la pérdida de capacidad funcional a partir de los 30 años ronda el 0,8 por ciento cada año. Además, en 1825, el inglés Benjamin Gompertz, experto de una compañía de seguros, descubrió que las probabilidades de morir se duplicaban cada ocho años a partir de los 30. ¿Cuáles son las causas de este declive natural? ¿Podría detenerse o incluso sería posible recuperar la plenitud funcional?

Exceptuando ciertos tipos celulares (especialmente las células nerviosas y algunas células musculares), el organismo de una persona no es hoy el mismo organismo que fue varios años atrás. Muchas de sus células han muerto y han sido reemplazadas por sus descendientes. Las nuevas células, sin embargo, no rejuvenecen al organismo. El envejecimiento no se manifiesta en las células individuales sino en las estirpes celulares. Entre los seres vivos más viejos se encuentran determinados árboles, como las secuoyas. Sin embargo, las células vivas de tales árboles no tienen más de 30 años, y la mayor parte del árbol está constituido por células muertas que no son necesarias para su supervivencia. Puesto que para determinar la edad no hay que considerar las células muertas, dichos árboles no tienen en realidad más de 30 años y son, por tanto, notablemente más jóvenes que las células nerviosas más viejas de muchos seres humanos.

En nuestro laboratorio del Hospital Pediátrico del Centro Médico de California septentrional estamos investigando el proceso de envejecimiento a nivel celular. Al cultivar in vitro fibroblastos humanos normales (las células estructurales de los tejidos de sostén del organismo), éstos se dividen repetidamente durante meses, poco a poco van dejando de dividirse y, finalmente, mueren. Tales observaciones sugieren que el proceso de envejecimiento de una célula normal es una propiedad innata de las células.

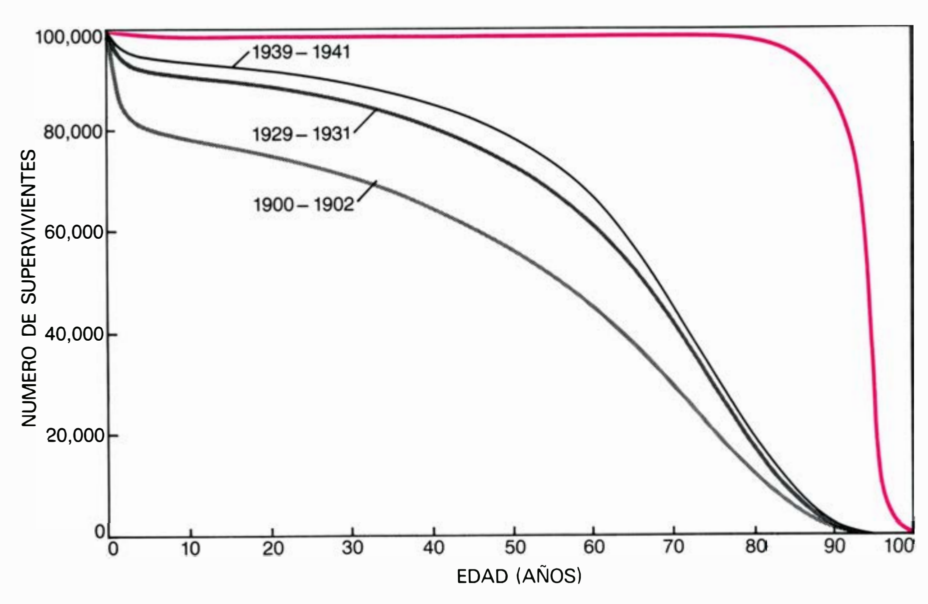
Nuestra técnica de cultivo celular es relativamente simple. Primero tratamos pequeñas cantidades de tejido pulmonar embrionario con enzimas digestivos que lo disocian en millones de células individuales. Las células se separan de los enzimas sedimentándolas por centrifugación, luego se introducen en botellas que contienen una disolución nutritiva y se incuban a temperatura corporal. Las células comienzan entonces a dividirse regularmente. Al cabo de una semana, la población celular alcanza un estado de confluencia, cubriendo toda la superficie del cultivo en forma de una capa unicelular. Una vez se ha alcanzado la confluencia, las células normales dejan de dividirse en virtud del fenómeno denominado inhibición por contacto.

Si se desean nuevas divisiones celulares, las células deben separarse de la botella que contiene el "cultivo madre" y distribuirse en números iguales en otras dos botellas con medio nutritivo fresco. Mediante este procedimiento, denominado subcultivo, se consigue, en el momento en que los cultivos hijos alcanzan la confluencia y dejan de dividirse, duplicar aproximadamente el número de células separadas del cultivo madre. El hecho de que la población celular se duplique no significa que cada célula se divida una vez; algunas células no se dividen y otras lo hacen varias veces. Si se preparan cuatro cultivos hijos a partir de un cultivo madre, el número total de células se duplicará dos veces, aunque para ello se requiere un período de incubación más largo. No obstante, con independencia de la dilución, la capacidad de división de los fibroblastos embrionarios humanos queda limitada a 50 duplicaciones de la población en un período de siete a nueve meses.

Cuando los fibroblastos normales agotan su capacidad de proliferación no dejan de dividirse súbitamente, sino que, a medida que la población celular se aproxima al límite de las 50 duplicaciones, se requiere un tiempo mayor para que los cultivos hijos alcancen la confluencia. El primer signo de senescencia próxima es la prolongación del tiempo empleado en alcanzar la confluencia, que pasa a ser de diez días en vez de una semana. El producto final es un subcultivo en el que nunca se alcanza la confluencia, con independencia de la dura-

FIBROBLASTOS "JOVENES" (células del tejido conectivo) de un embrión humano. La micrografía muestra cómo estas células recubren la superficie de una placa. Las celulas fusiformes se fijaron y tiñeron durante un período de proliferación activa. Los núcleos muestran un color rojizo.

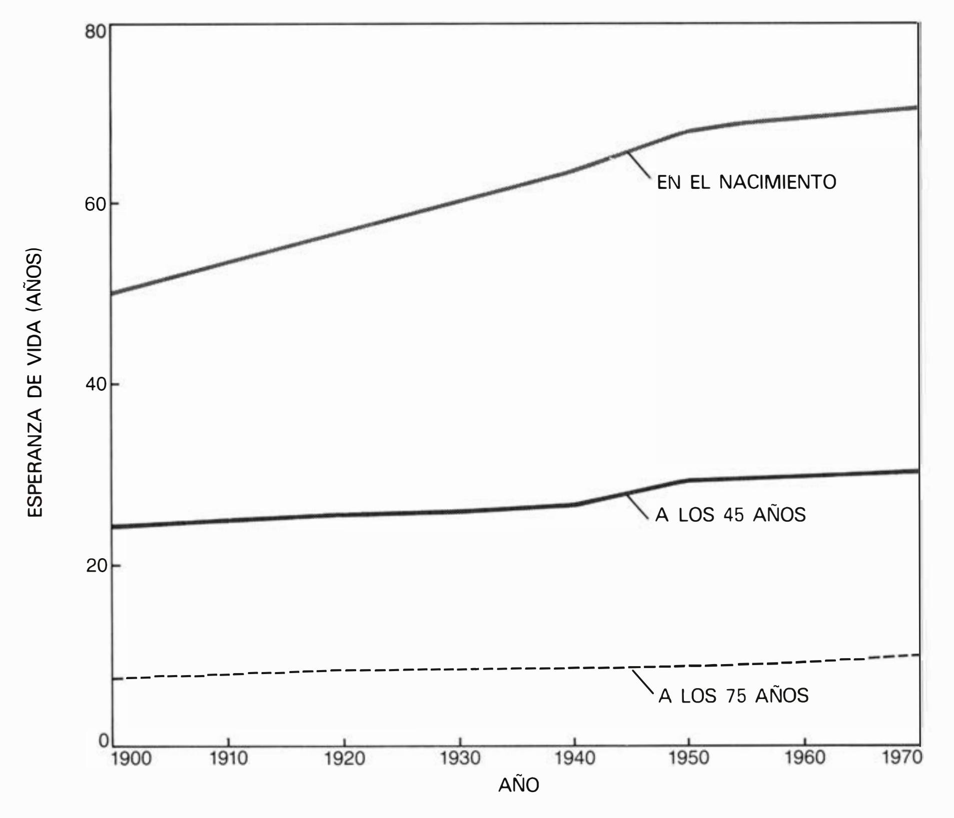
FIBROBLASTOS ENVEJECIDOS teñidos durante la pérdida de su capacidad de división. Dicha pérdida ocurre aproximadamente cuando han tenido lugar 50 duplicaciones de la población celular. Las células experimentan distintos cambios degenerativos y luego mueren. Esta observación sugiere que el envejecimiento es una propiedad innata de las células normales. Las micrografías fueron realizadas en el laboratorio del autor.



CURVAS DE SUPERVIVENCIA HUMANA para personas de raza blanca en tres períodos de la historia de los Estados Unidos. Las curvas muestran que, con la mejora de las condiciones de vida y los avances de la medicina, ocurren menos muertes en los primeros años de vida. La eliminación de todas las causas de muerte atribuibles a enfermedad o accidente daría lugar a una curva hipotética (*que es la curva en color; la primera empezando por arriba*); las personas morirían únicamente por los déficits fisiológicos relacionados con el envejecimiento. Los datos fueron recogidos por Alex Comfort.

ción del cultivo o de la frecuencia de la renovación del medio. Finalmente, las células experimentan distintos cambios degenerativos y mueren.

En un principio atribuimos la muerte de los cultivos a errores experimentales; la división celular, como se sabe, resulta adversamente afectada por defectos en la preparación del medio de cultivo, por limpieza inadecuada del material de vidrio, por contaminación vírica y por otros artefactos técnicos. Intensos estudios sugirieron, no obstante, que había otras razones. El experimento más demostrativo de que la muerte celular es una propiedad innata de las células normales lo realizamos en colaboración con Paul S. Moorhead. Aprovechando el hecho de que las cepas de fibroblastos humanos, derivadas de embriones masculinos y femeninos, pueden diferenciarse comparando sus cromosomas sexuales, mezclamos células femeninas con una edad de 10 duplicaciones de la población con igual número de células masculinas con una edad correspondiente a 40 duplicaciones de la población. Se utilizaron cultivos sin mezclar de cada cepa, como control; los tres cultivos fueron sometidos a subcultivo de la forma habitual. Tras 20 duplicaciones de la población, en los cultivos mixtos sólo logramos distinguir células femeninas. Los cultivos control de células femeninas seguían dividiéndose, pero las células control masculinas habían muerto. Estos resultados apoyan la idea de que la muerte de las células masculinas más viejas era atribuible a una propiedad innata de las propias células y no a errores de la técnica, pues en otro caso sería difícil explicar por qué únicamente sobrevivían las células femeninas del cultivo mixto.



ESPERANZA DE VIDA (*en ordenadas*) para personas de raza blanca en los Estados Unidos de América en el nacimiento y a los 45 y 75 años de edad a partir de 1900. El incremento de alrededor de 24 años en la esperanza de vida desde 1900 ha tenido lugar principalmente en los grupos jóvenes.

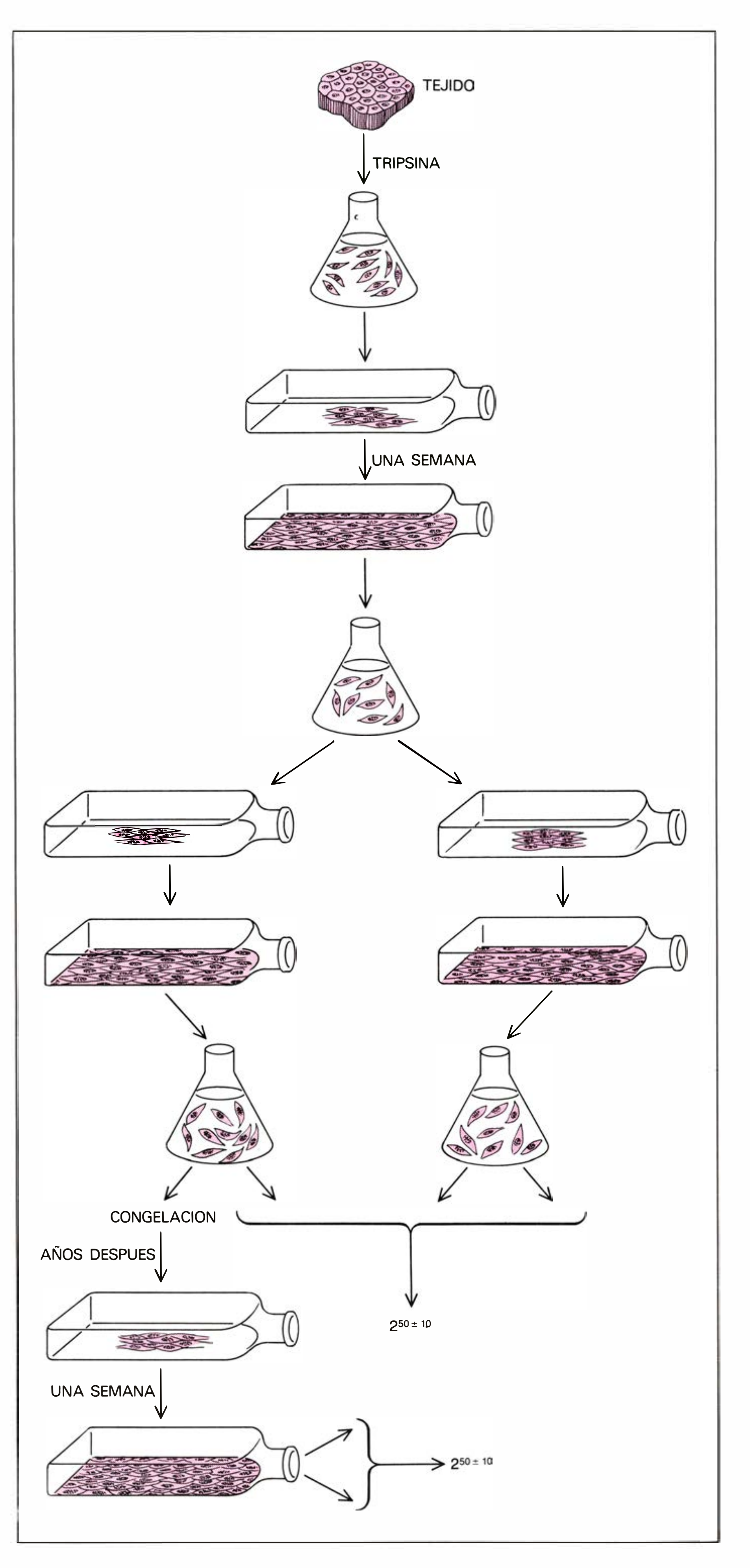
Otra observación sugirió que las cepas de fibroblastos normales tenían un período vital limitado. Cuando se almacenaban los fibroblastos procedentes de embriones humanos debajo de cero grados Celsius, manifestaban una sorprendente "memoria". Si habían sido congelados en el momento de alcanzarse la vigésima duplicación de la población, después de descongelados experimentaban 30 duplicaciones más y luego se detenían. Si se congelaban en la décima duplicación de la población celular, y luego se descongelaban, dichas células experimentaban 40 duplicaciones más. Es decir, en todo caso, la suma de duplicaciones antes y después de la congelación totalizaba 50. Era como si las células poseyeran un contador o programador de divisiones. ¡Una de nuestras cepas celulares humanas designada WI-38 conservaba aún esta memoria tras más de 13 años de preservación en nitrógeno líquido!

La duración limitada de la vida de células humanas en medio de cultivo ha sido observada en fibroblastos y otros tipos celulares normales, derivados de distintos tejidos tales como la piel, el cerebro, el hígado y la fibra muscular lisa. No se conoce todavía ninguna excepción a la regla general de que las células normales poseen una capacidad limitada para dividirse. Las cepas de células humanas en medio de cultivo pueden transformarse, no obstante, en líneas celulares "inmortales" al ser tratadas con el virus SV 40, productor de cáncer. Se conocen alrededor de unas 600 líneas

celulares inmortales o transformadas, muchas de las cuales se desarrollan espontáneamente en los cultivos celulares. La línea celular humana más famosa, llamada HeLa, fue obtenida en 1952 a partir de un cultivo celular del cérvix uterino. Ha sido mantenida en cultivo desde entonces. Las células transformadas son claramente anormales en características tales como el número y la forma de sus cromosomas, sus propiedades químicas y sus características tintoriales. Es más, muchas de las células transformadas dan lugar a tumores al ser inyectadas en animales de laboratorio, mientras que las células normales no producen tal efecto. Se llega, por tanto, a la paradójica conclusión de que las células animales deben adquirir algunas de las propiedades de las células cancerosas para poseer una capacidad ilimitada de dividirse.

Algunos descubrimientos posteriores nos sugirieron que el tiempo vital limitado de las células normales en medio de cultivo podría representar una forma de envejecimiento a nivel celular. Cultivamos fibroblastos obtenidos de personas de edad avanzada para determinar si las células más viejas experimentaban un menor número de divisiones en el medio de cultivo. Puesto que nuestros estudios iniciales se habían realizado a partir de tejido pulmonar de embriones humanos, utilizamos fibroblastos pulmonares procedentes de autopsias de ocho personas adultas de edades comprendidas entre los 20 y los 87 años.

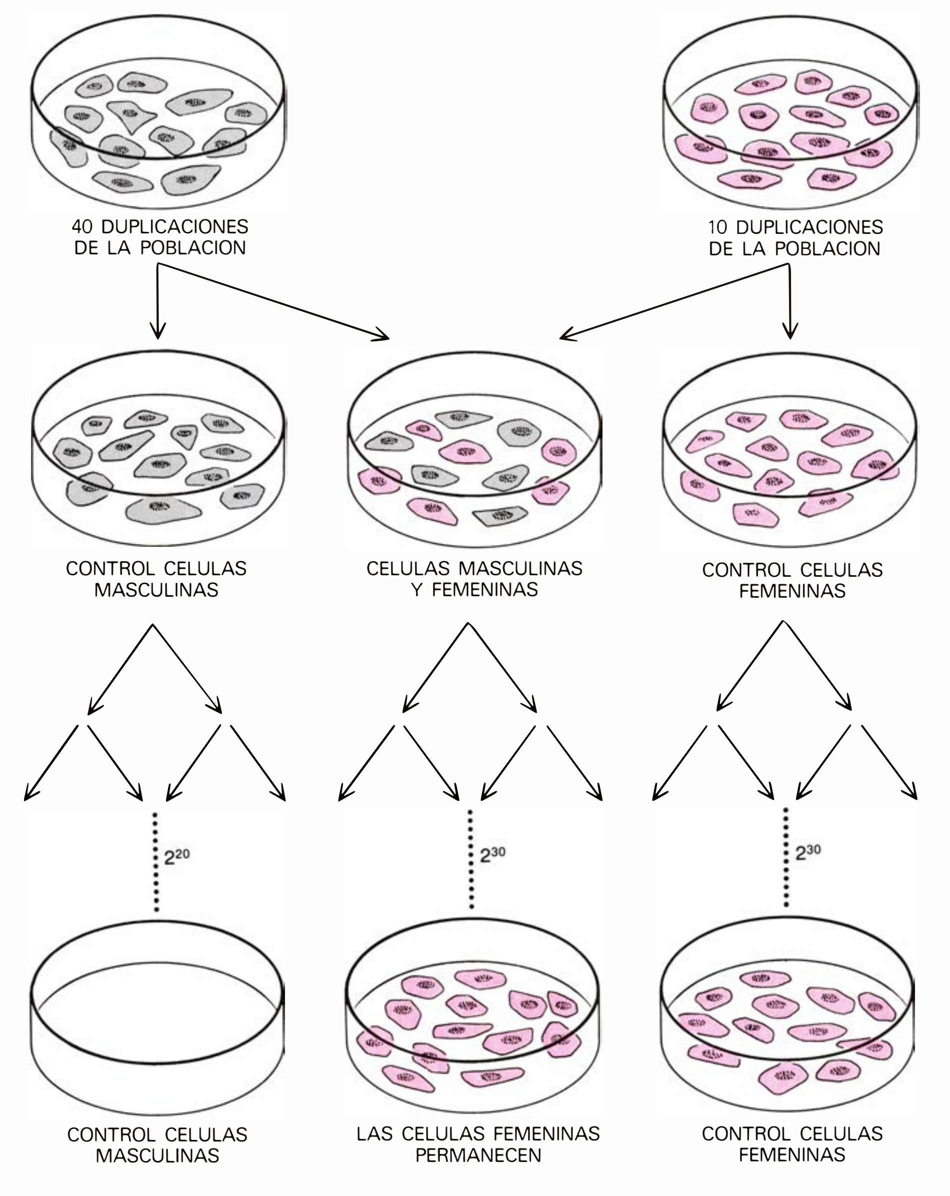
Observamos que los fibroblastos viejos se duplicaban entre 14 y 29 veces en medio de cultivo, sin que existiera una clara correlación entre el número de divisiones y la edad del dador. Hubo, no obstante, una reducción considerable en el número de duplicaciones en relación con las 50 que, de forma invariable, se observaban en los fibroblastos embrio-



CULTIVO DE CELULAS HUMANAS normales. La técnica utilizada en el laboratorio del autor consiste en la disociación del tejido en células individuales mediante la utilización del enzima digestivo tripsina. Las células se transfieren a continuación a una botella aplanada, donde se multiplican hasta cubrir la superficie del recipiente. Luego dejan de multiplicarse. La población celular se divide entonces en dos mitades iguales, cada una de las cuales vuelve a cultivarse. Este proceso puede repetirse únicamente unas 50 veces con células humanas fetales; no puede repetirse tantas veces utilizando células de donantes adultos. Las células cultivadas pueden almacenarse congeladas en nitrógeno líquido. Si al cabo de varios años las células vuelven a cultivarse, aún "recuerdan" el nivel de duplicaciones al que fueron congeladas y continúan dividiéndose hasta alcanzar el número de duplicaciones posibles.

narios humanos. Más recientemente George M. Martin y colaboradores, de la Facultad de Medicina de la Universidad de Washington, Samuel Goldstein, de la Universidad McMaster en Canadá, y Edward Schneider y Youji Mitsui, del Centro de Investigación Gerontológica de Baltimore, observaron en estudios realizados en cultivos de fibroblastos obtenidos a partir de biopsias de piel humana, que el número de duplicaciones celulares es inversamente proporcional a la edad del dador. Resultados similares fueron obtenidos con células hepáticas humanas por Y. Lebuilly, en Francia, y con células musculares por Edwin L. Bierman, de la Facultad de Medicina de la Universidad de Washington.

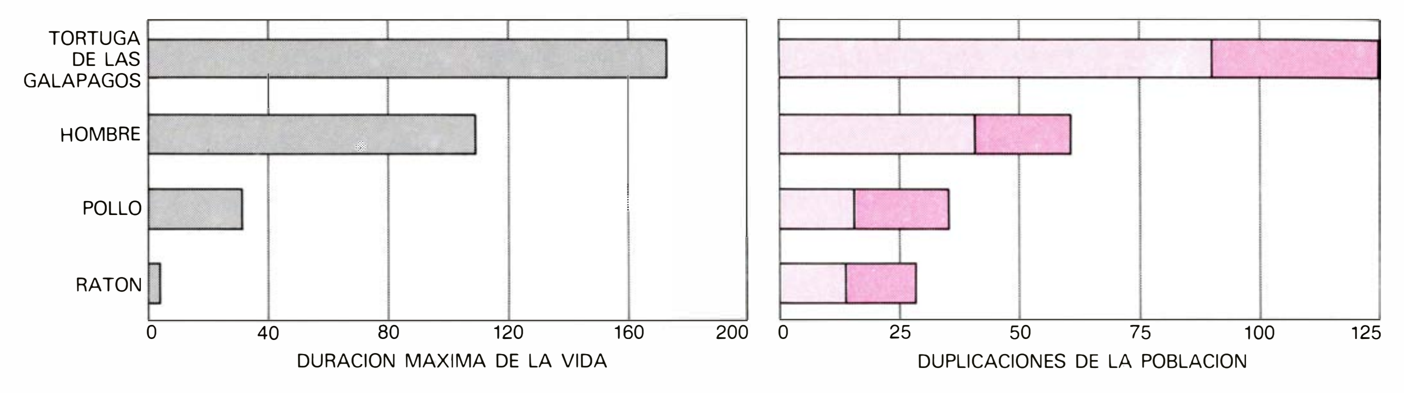
El envejecimiento es una característica de todos los animales que alcanzan un tamaño fijo con la madurez. Algunos peces continúan creciendo a lo largo de toda su vida y no envejecen en el sentido habitual del término. Resulta interesante plantearse si las células normales de diferentes animales, al igual que las humanas, envejecen en cultivo celular. Las variaciones en la duración de la vida en diferentes especies son mucho mayores que las existentes entre individuos de la misma especie. La mosca de la fruta es vieja a los 40 días, un ratón a los tres años, un caballo a los 30, el hombre a los 100 y ciertas tortugas a los 150. Aunque sólo se ha examinado algunas especies, los datos disponibles sugieren una relación directa entre la duración de la vida en una determinada especie y la capacidad de sus células de dividirse en medio de cultivo. Otra relación interesante ha sido puesta de manifiesto por George A. Sacher, Jr., del laboratorio Nacional de Argonne. Dicho investigador observó una fuerte correlación positiva entre la duración de la vida y el cociente de dividir el peso cerebral por el peso total del organismo. Esta correlación fue observada en 85 especies animales de diferentes tamaños, desde el ratón al elefante. La observación indica que a lo largo de la evolución de los animales vertebrados puede haber existido una importante relación entre el incremento de la duración de la vida y el incremento del tamaño cerebral.



EL LIMITE DE DUPLICACION de células normales en medio de cultivo se demostró preparando un cultivo mixto de células fetales femeninas y masculinas que eran respectivamente "jóvenes" y "viejas". Los cultivos separados de cada tipo sirvieron de control. Después de 30 duplicaciones de la población, sólo las células femeninas permanecían en el cultivo mixto y únicamente las células control femeninas seguían dividiéndose. Las células masculinas habían agotado su potencial de duplicación y habían dejado de dividirse varias semanas antes de que la mezcla fuera examinada. Es improbable que errores de tipo técnico o virus contaminantes pudieran explicar el que únicamente las células masculinas de los cultivos mixtos hubieran muerto. La explicación que resultaba, con mucho, la más verosímil era aquella que afirmaba que la pérdida de la capacidad normal de división se debía a un control intracelular.

Un segundo enfoque para estudiar el envejecimiento celular consiste en realizar trasplantes en serie de células marcadas en animales consanguíneos antes de que éstos envejezcan. Con ello se evita la muerte de las células marcadas cuando el animal huésped muere. Tales experimentos in vivo han completado los realizados en cultivos celulares. Los resultados de distintos laboratorios sobre trasplantes en serie de tejido de glándula mamaria, piel y células sanguíneas inmaduras en ratas, indicaron que las células trasplantadas envejecían y morían igual que lo hacían en cultivos celulares. En algunos casos, sin embargo, el tiempo transcurrido hasta la muerte de las células era mayor que la duración de la vida de un animal de dicha especie. Esta observación puede explicarse por el hecho de que las células trasplantadas no crecen tan rápidamente como lo harían en medio de cultivo, sino que la mayoría del tiempo permanecen en estado de reposo. Si crecieran al mismo ritmo que el cultivo celular, el trasplante alcanzaría en pocos meses un tamaño varias veces superior al del huésped.

Un fenómeno similar ha sido observado al cultivar células normales a temperatura ambiente en lugar de hacerlo a la temperatura corporal; las células se dividían lentamente y degeneraban muchos meses después de la muerte de cultivos hermanos mantenidos a la temperatura corporal. Aunque el tiempo de supervivencia aumentaba, el número total de duplicaciones de la población celular no excedía, en las células incubadas a temperatura ambiente, al correspondiente a las células incubadas a temperatura corporal. En contraste con lo anterior y al igual que ocurría en cultivos celulares, donde las células transformadas eran inmortales, las células cancerosas se replican indefinidamente al transplantarlas de unos animales a otros.



CAPACIDAD REPLICATIVA de fibroblastos fetales (de animal recién nacido en el caso de la tortuga) en medio de cultivo. Dicha capacidad es proporcional a la duración de la vida de las distintas especies. La limitada capacidad de duplicación in vitro se correlaciona con el envejecimiento.

Para localizar e interpretar el mecanismo responsable de la limitada capacidad de replicación de las células normales, utilizamos, trabajando en colaboración con Woodring Wright, un enfoque experimental distinto. Al tratar las células cultivadas con citochalasina *B*, las células expulsan los núcleos y estos pueden separarse de los restos celulares por centrifugación. De esta forma pueden obtenerse millones de células desprovistas de sus núcleos. Dichas células se denominan citoplastos y permanecen viables durante varios días. A lo largo de ese intervalo pueden fusionarse con células normales. Para determinar si el programador que dicta la capacidad replicativa está localizado en el núcleo o fuera de él, en el citoplasma, fusionamos citoplastos derivados de células jóvenes con células viejas y viceversa. Observamos que el citoplasma tenía poco efecto en la rapidez de envejecimiento; ello sugería que el control radi-

DURACION DE LA VIDA (AÑOS)
100
10
1
.1
1
10
100
1,000
10,000
PESO DEL CEREBRO (GRAMOS)

O PRIMATES
+ ELEFANTE
● ARTIODACTILOS
■ PERISODACTILOS
▲ ROEDORES
O MARSUPIALES
+ MURCIELAGOS
● CARNIVOROS
■ INSECTIVOROS
▲ CETACEOS

EL TAMAÑO CEREBRAL parece también estar relacionado con la duración de la vida de varias especies animales, sugiriendo que existe una conexión entre ambos parámetros, que debió aparecer durante la evolución de los animales vertebrados. Como muestra esta representación logarítmica, hay una correlación significativa con una pendiente de alrededor de 0,25. La dispersión de los datos se debe a las variaciones de la duración de la vida para un determinado peso cerebral y a la estratificación entre órdenes de animales. Al tomar en consideración otras variables (tales como el peso corporal, el metabolismo y la temperatura corporal), la dispersión alrededor de la línea se reduce considerablemente. Los datos mostrados en la gráfica fueron recogidos por George A. Sacher, Jr., del laboratorio de Argonne.

caba probablemente en el núcleo. Recientemente se han desarrollado técnicas para insertar núcleos celulares en citoplastos. En colaboración con Audrey Muggleton-Harris hemos reconstituido células humanas normales viables, a partir de núcleos aislados jóvenes y de citoplastos viejos, así como partiendo de núcleos viejos y citoplastos jóvenes. Los resultados de los primeros ensayos parecen indicar que el programador que determina la capacidad de proseguir las divisiones celulares se encuentra realmente en el núcleo.

En lugar de disociar el tejido en sus células individuales, resulta también posible cultivar pequeños fragmentos de un tejido animal intacto directamente sobre vidrio o plástico. Ya en los años veinte se descubrió, en un pequeño fragmento de tejido de embrión de pollo colocado en cultivo celular, que el tiempo que tardaban los fibroblastos en salir (migrar) de los límites del tejido cultivado aumentaba con la edad del embrión. Este lapso de tiempo se denomina período de latencia.

En los años sesenta Milena Soukupová y Emma Holecková, de la Universidad Carolina de Praga, observaron un incremento similar en el período de latencia en relación con la edad de las ratas en cultivos de fragmentos de corazón, hígado y riñón. En 1970, Roy L. Walford, Jr. y Harold Waters, de la Facultad de Medicina de la Universidad de California en Los Angeles, obtuvieron 20 cultivos de fragmentos de tejido de donantes humanos con edades comprendidas entre los pocos minutos y los 80 años, y observaron que el período de latencia aumentaba con la edad (aunque aumentaban también las variaciones individuales del período).

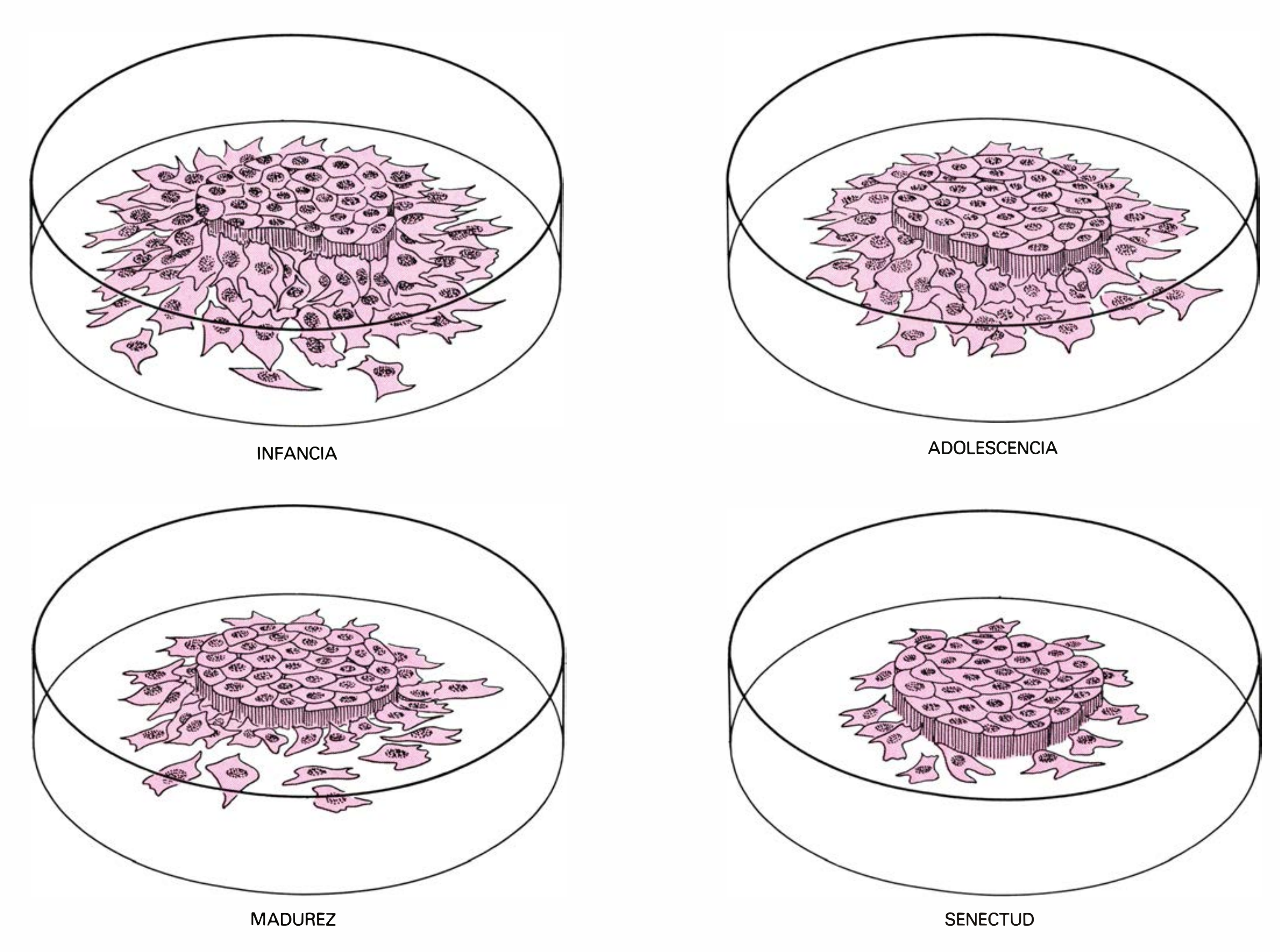
A partir del hecho comprobado de que las células normales no son inmortales, sino que tienen una capacidad limitada para dividirse, resulta importante determinar el impacto de este descubrimiento en la comprensión del proceso del envejecimiento. Los datos actuales indican la existencia de algún tipo de programación en las células normales que controlaría su capacidad funcional y reproductora. No todas las células del organismo, sin embargo, tienen la capacidad de los fibroblastos para una replicación rápida. Los tejidos en los que las células experimentan replicación rápidamente son la piel, los tejidos formadores de sangre y la capa celular que recubre el intestino. Algunas células más especializadas, como las células nerviosas, las células endocrinas, las células musculares, las células sensoriales y algunas células del sistema inmunitario experimentan pocas divisiones o no se dividen en absoluto una vez han alcanzado la madurez. Los gerontólogos están de acuerdo en general en el hecho de que los cambios más importantes que ocurren con la edad tienen lugar en las células altamente especializadas y no en las células que experimentan rápidas divisiones. De hecho, parece probable que el animal envejezca y muera por otras razones antes de que se haya alcanzado el punto final de la replicación de los fibroblastos y de otras células que se dividen rápidamente.

No obstante, el mecanismo que limita la división celular en las células que proliferan rápidamente es probablemente el mismo que limita la capacidad funcional de las células especializadas que se dividen más lentamente o que no experimentan división. En la pasada década se detectó una gran variedad de cambios fisiológicos y morfológicos en cultivos de fibroblastos humanos mucho antes de que pierdan su capacidad de dividirse. Entre ellos pueden citarse los cambios en la utilización de las sustancias nutritivas, en la expresión y reparación del material genético, en las vías metabólicas de la célula, en la actividad de varios enzimas cruciales y en la morfología de la superficie celular y de los orgánulos celulares. Muchos de estos cambios son idénticos a los que ocurren en los seres humanos con el envejecimiento. Una explicación del mecanismo a través del cual los fibroblastos humanos pierden su capacidad de dividirse, puede arrojar luz sobre los factores determinantes de las pérdidas funcionales por envejecimiento en células más especializadas.

¿Cuáles son, pues, los mecanismos subyacentes a los cambios inducidos por el envejecimiento en la mayoría de las células? Muchos gerontólogos creen que la respuesta puede encontrarse en el mensaje genético. Su razonamiento es el siguiente: si el complejo desarrollo desde el huevo fertilizado hasta la madurez sexual está orquestado por el aparato genético, los cambios produci-

ESTIRPE CELULAR JOVEN (10 DIVISIONES)
ESTIRPE CELULAR VIEJA (30 DIVISIONES)
SOBREVIVE 40 DIVISIONES MAS
SOBREVIVE 20 DIVISIONES MAS

LOS EXPERIMENTOS DE FUSION CELULAR han ayudado a determinar la localización del programador que controla el envejecimiento celular. Ello fue posible gracias al desarrollo de técnicas de aislamiento del núcleo celular y de reinserción en otras células. Las células desprovistas de núcleo, denominadas citoplastos, pueden sobrevivir durante un corto período de tiempo. Los núcleos de células jóvenes (10 duplicaciones de población) fueron insertados en citoplastos de células viejas (30 duplicaciones de población) y viceversa. Los resultados preliminares sugirieron que el programador que controla el número de divisiones celulares se encuentra en el núcleo y no en el citoplasma que rodea el núcleo.



TIEMPO DE MIGRACION de fibroblastos que crecen a partir de un pequeño fragmento de tejido en medio de cultivo. Varía inversamente con la edad del organismo de donde proceden. Este hecho ha sido demostrado con fibroblastos de pollo, rata y seres humanos, obtenidos de diferentes tejidos.

dos por el envejecimiento serían controlados también, probablemente, por los genes. Tres hipótesis generales basadas en las propiedades de las moléculas portadoras de información (ADN y ARN) se consideran actualmente como las explicaciones más plausibles del proceso del envejecimiento.

La primera hipótesis, formulada por Zhores Medvedev, del Medical Research Council en Londres, y desarrollada ulteriormente por Leslie Orgel, del Instituto Salk, propone que, con el tiempo, la información transmitida en los procesos de transcripción y traducción del mensaje genético, desde el ADN al ARN y a los enzimas y otras moléculas proteicas, podría estar sujeta a un número progresivamente mayor de errores. Estos errores darían lugar a moléculas enzimáticas defectuosas, y conducirían a un declive de la capacidad funcional de las células. La situación sería análoga a la que se produce por funcionamiento defectuoso de una pieza en una máquina automática. El instrumento daría lugar, a su vez, a partes defectuosas que, al ser ensambladas en el producto final, reducirían su eficiencia o impedirían su función. Aunque tal hipótesis sobre la existencia de los "errores catastróficos" ha sido sometida a prueba en varios laboratorios estudiando la síntesis proteica en células en proceso de envejecimiento, los resultados no han proporcionado pruebas adecuadas para sostener dicha hipótesis.

La acumulación de errores en un sistema biológico puede evitarse en parte por los conocidos procesos de reparación, pero los propios sistemas de reparación no actúan a la perfección ni de forma indefinida. Ronald W. Hart y Richard B. Setlow, del Laboratorio Nacional de Oak Ridge, observaron que fibroblastos de la piel obtenidos de diferentes animales, desde la musaraña al elefante y al hombre, eran capaces en medio de cultivo de reparar las alteraciones de su ADN causadas por radiaciones ultravioleta. La capacidad de reparación variaba en proporción directa a la duración de la vida de las especies estudiadas. Por ejemplo, los seres humanos viven alrededor del doble que los chimpancés, y la velocidad de reparación del ADN en aquéllos viene a ser también aproximadamente el doble. Más recientemente, estos y otros investigadores han confirmado los estudios anteriores y han observado que la capacidad para la reparación del ADN en cultivos de células normales disminuye a medida que se aproximan al límite de su capacidad de replicación.

Estos procesos pueden compararse con lo que sucede en un automóvil o en una lavadora automática. Las dos máquinas tienen lo que los ingenieros denominan "duración media sin fallo mecánico". Sin reparaciones pueden funcionar normalmente durante algunos años, y con reparaciones perfectas pueden durar indefinidamente mientras se disponga de piezas de recambio. Si las reparaciones no son perfectas o faltan las piezas adecuadas, la máquina irá a parar finalmente al montón de la chatarra. Las piezas de recambio quirúrgicas para seres humanos están limitadas a unos pocos órganos, y tales piezas son en sí mismas imperfectas. Aun cuando los trasplantes cerebrales fueran posi-

bles, el esfuerzo sería en vano. La inmortalidad, después de todo, presupone la continuidad de la propia identidad, algo imposible con un cerebro ajeno.

Medvedev ha propuesto también la segunda hipótesis sobre las bases genéticas del envejecimiento. Es sabido que sólo alrededor de un 0,4 por ciento de la información del ADN contenido en el núcleo celular es utilizada por una determinada célula en su período vital. Es más, muchos de los genes de la molécula del ADN están repetidos en secuencias idénticas, por lo que el mensaje genético resulta altamente redundante. Medvedev ha propuesto que las secuencias repetidas estarían normalmente reprimidas, pero cuando un gen activo fuera dañado extensamente, sería reemplazado por un gen idéntico de reserva. La redundancia del ADN podría, por lo tanto, proporcionar un mecanismo protector frente a la vulnerabilidad intrínseca del sistema, causada por accidentes moleculares. A través de este mecanismo se prolongaría el tiempo de supervivencia al evitar el acúmulo de un número suficiente de errores capaz de alterar el mensaje genético. Al final, sin embargo, todos los genes repetidos serían utilizados, los errores se acumularían y las deficiencias fisiológicas determinantes del envejecimiento irían apareciendo. Esta hipótesis permite hacer la predicción de que las especies de larga vida deben tener más ADN redundante que las especies de vida corta.

La tercera hipótesis genética del envejecimiento propone que los cambios producidos por la edad son simplemente una continuación de las señales genéticas normales que regulan el desarrollo de un animal desde el momento de la concepción a la madurez sexual. Supone incluso la existencia de "genes de envejecimiento" que frenarían o detendrían vías bioquímicas de forma secuencial y conducirían a una expresión programada de los cambios propios del envejecimiento. La coloración gris del cabello, la menopausia y la disminución de las facultades atléticas son unos pocos de entre los cientos de fenómenos asociados al envejecimiento. Ninguno de ellos es considerado como una enfermedad, pero tales cambios disminuyen la capacidad funcional de la célula e incrementan la vulnerabilidad a las enfermedades. Dichos fenómenos programados genéticamente variarían en cuanto al tiempo de su expresión en diferentes tipos celulares. La raíz del envejecimiento podría encontrarse en deficiencias de tipos celulares clave cuya velocidad de envejecimiento es más rápido y tiene mayores consecuencias.

La función de los hipotéticos genes de envejecimiento podría ser análoga al declinar funcional normal y a la muerte de las células que tiene lugar a escala masiva durante el desarrollo del embrión. En los animales vertebrados, el desarrollo de una extremidad, por ejemplo, supone no sólo la formación de millones de células sino también la muerte y la resorción de millones de otras células. El destino de estas últimas está determinado por "programadores de muerte" que operan en el momento oportuno. Puede suponerse que el mismo proceso continúa a lo largo de la vida, operando a diferentes velocidades en distintos tejidos y finalmente produce los cambios que conducen al envejecimiento y que aumentan la susceptibilidad a las enfermedades. Los defensores de un tal programa genéticamente predeterminado para el envejecimiento argumentan que el éxito evolutivo de las especies depende sólo de la capacidad de sus miembros para vivir el tiempo necesario para procrear a sus hijos y criarlos; lo que sucede luego es irrelevante para la supervivencia de la especie.

Las tres hipótesis sobre el envejecimiento pueden aplicarse tanto a células que crecen en cultivo como a las células que envejecen en el organismo intacto. Las tres hipótesis no se excluyen mutuamente y podrían operar de un modo simultáneo. Dos líneas celulares escapan del proceso de envejecimiento y muerte: las células cancerosas y las células germinales (el óvulo, el espermatozoide y sus inmediatos precursores). Resulta interesante especular si las células cancerosas y las células germinales podrían librarse del envejecimiento a través de un mecanismo común. De hecho, la forma en que la información genética es intercambiada entre los virus productores de cáncer y las células animales podría ser similar a la forma en que las cartas genéticas vuelven a barajarse cuando el espermatozoide se fusiona con el óvulo. El intercambio de información genética entre dos células normales adyacentes en el organismo (mediada por virus, agentes carcinogénicos o radiaciones ionizantes) podría originar una célula cancerosa. Al barajar la información genética en la génesis del óvulo y el espermatozoide o en el proceso de fusión de estas células volvería a ponerse en marcha el mecanismo que controla el número de divisiones celulares. De esta forma, aunque los distintos miembros de una especie estarían programados para morir, la especie sobreviviría. Un ser humano sería, pues, el camino a través del cual las células germinales producirían nuevas células germinales inmortales.

Reservas uraníferas mundiales

Un estudio de la estimación de reservas de minerales, basado en informes mineros de los Estados Unidos, muestra que el abastecimiento de uranio no será ningún factor limitante para el desarrollo de la energía nuclear

Kenneth S. Deffeyes e Ian D. MacGregor

La estimación de los recursos mundiales de minerales es un empeño bastante reciente en geología. Nuestro interés en el tema se remonta a 1972, cuando un grupo de analistas del Instituto de Tecnología de Massachusetts publicó el libro titulado *Los límites del crecimiento*. El libro se basa en un trabajo que hicieron intentando estudiar el futuro mediante la confección de un modelo cuantitativo de la economía mundial y que, con un ordenador, proyectaban luego a los tiempos venideros. A causa del agotamiento de los recursos, el modelo predice un desastre hacia el año 2000. Como geólogos estamos interesados en todo lo que, en este contexto, pueda relacionarse con el término recursos. Una lectura más cuidadosa nos revela que el término "recursos" se empleó realmente con el mismo sentido con que los geólogos usamos la palabra "reservas", esto es, para indicar la cantidad de material que ya está descubierto y que es recuperable, con provecho, en las condiciones económicas actuales. Para sorpresa nuestra no pudimos encontrar ninguna estimación cuantitativa sobre los yacimientos de minerales sin descubrir, o sobre los yacimientos ya conocidos y que en el futuro pudiesen pasar a una categoría económicamente rentable. Ante un estudio del tipo representado por *Los límites del crecimiento* se podría pensar en dar un pronóstico mejor; habría que intentar estimar la disponibilidad total de los recursos minerales.

Para el caso del uranio era muy urgente una estimación semejante de los recursos. Por ejemplo, la justificación económica inicial del reactor rápido regenerador, que se había propuesto casi a la vez que apareció el estudio *Los límites del crecimiento*, suponía que la curva de suministro de uranio (la curva que relaciona el precio del mineral con la cantidad recuperable a ese precio) seguía una trayectoria ascendente uniforme [*véase la ilustración de la página 37*]. No obstante, parecía muy posible en aquel tiempo que un pequeño incremento de precio permitiría teóricamente que todos los yacimientos de uranio de una categoría geológica dada fuesen susceptibles de descubrimiento y explotación. Luego sería necesaria una gran alza de precios antes de que ninguna otra categoría se convirtiese en explotable. En ese caso, la curva de suministro ascendería primero y luego se aplanaría antes de empezar a ascender de nuevo.

Si tal curva describiese exactamente la situación uranio-suministro, la elevación del precio a niveles intermedios equivaldría, o bien a una compensación en metálico a los propietarios de los yacimientos conocidos, o bien a un incentivo para gastar dinero, esfuerzos y energía en la búsqueda de yacimientos de ley media no inventariados. (Sabemos que algunos geólogos especialistas en petróleo opinan que los precios superiores a siete dólares por barril corresponden ya a los del bache representado por la parte plana de la curva coloreada de aprovisionamiento que hay en la figura de la página 37). A pesar de la necesidad evidente de tener estimaciones de recursos que merezcan confianza, tanto en lo que se refiere al combustible nuclear como a los combustibles fósiles, metales u otros materiales de importancia vital, no había en 1972, ni lo hay todavía hoy, ningún método de evaluación que sea aceptable de forma general. En realidad, para la mayoría de los metales todavía no se ha hecho acopio de los datos geológicos primarios en que tendría que basarse cualquier método rudimentario de estimación de recursos.

De los pocos métodos de estimación de recursos que se han propuesto en los últimos años, la mayoría son muy específicos y limitativos para un material dado. Vicent E. McKelvey, del Servicio Geológico de los Estados Unidos, propuso en 1972 un modelo que pretendía fuese aplicable a todo tipo de recursos. Estableció una malla de categorías con dos entradas, que estaba basada en la magnitud de los conocimientos geológicos (yacimientos conocidos, yacimientos inferidos y yacimientos probables) y en su categoría económica actual (rentable, marginal y antieconómica). La única forma de estimar la cantidad de material en cada una de las categorías de McKelvey es preguntar a varios expertos que hagan suposiciones de las cantidades de material en cada categoría, sumar luego las estimaciones y hallar la media dividiendo la suma por el número de expertos. Aunque el Servicio Geológico fue más allá, tratando de cifrar la incertidumbre de las estimaciones de los expertos, para el petróleo todavía no descubierto, es evidente que resulta muy difícil basar un plan de acción estatal sobre tales datos subjetivos de partida.

En la década de los cincuenta, M. King Hubbert, de la Shell Development Company, abrió un camino de acceso, completamente diferente, para la estimación de recursos. Analizó la historia de la producción de petróleo en los Estados Unidos y ajustó una curva simétrica en forma de campana, con la capacidad anual de producción local de petróleo. En los tiempos de la primera estimación de Hubbert para la producción de petróleo se estaba todavía en una etapa embrionaria en los Estados Unidos y se fio de la opinión de los expertos para definir el área total cubierta por la curva en forma de campana. En la década de los sesenta, sin embargo, había recogido, directamente, suficientes cifras de la producción anual para poder deducir una estimación de los recursos de petróleo en los Estados Unidos y no tener que acudir a la opinión de los expertos. La realidad histórica ha sido favorable a la estimación inicial de Hubbert: la marcha reciente de la producción petrolífera interna sigue bastante bien su predicción de 1957 [*véase la ilustración superior de la página 38*].

A lo largo de la década de los sesenta, los métodos de Hubbert se criticaron duramente, debido a que empleaba como variable independiente el tiempo, en lu-

gar de usar alguna magnitud que representase el esfuerzo dedicado a la exploración del petróleo. En 1967 introdujo una segunda consideración en la que comparaba los metros de sondeos de exploración perforados en la búsqueda de petróleo y la cantidad de petróleo descubierto. Encontró que había una disminución exponencial del número de barriles de petróleo descubiertos por cada metro de sondeo perforado a medida que progresaba la exploración. Como era de esperar, la estimación de Hubbert para los recursos petrolíferos de los Estados Unidos deducida a partir del área limitada por la curva "esfuerzo de exploración" fue muy parecida a la estimación que había deducido a partir del área cubierta por la curva de campana de producción anual.

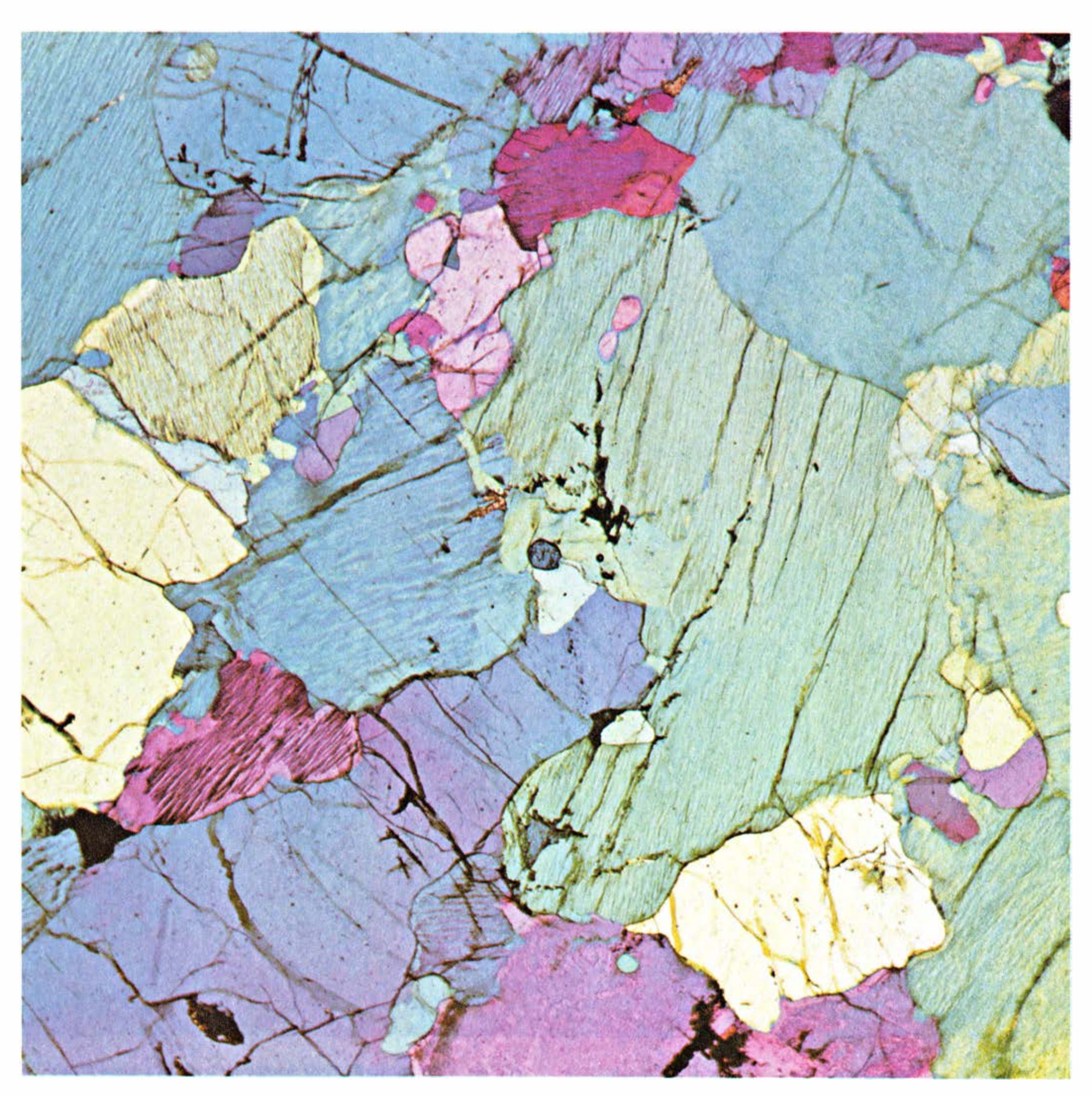
Cabía la pregunta de si los métodos de Hubbert podrían servir de base para hacer estimaciones de recursos de otros minerales y, en particular, de uranio. No había ninguna esperanza en la aplicación al uranio del primer esquema basado en la curva campaniforme de producción anual. La historia de la minería del uranio está caracterizada por ciclos de "auges y fracasos", y no habría ninguna previsión razonable deducida de tal dependencia con el tiempo. Hay, sin embargo, otra curva gaussiana que es bien conocida de los geólogos. En 1954 Louis H. Ahrens, de la Universi-



EL URANIO ES TRANSPORTADO desde el manto de la tierra hasta la parte superior de la corteza continental. En los bloques de granito que han aflorado queda como un elemento-traza. La imagen, en color falso, tomada por el ERTS-1 muestra uno de estos bloques, que pertenece a una región desértica de Australia. Se fotografío desde una altitud de 912 kilómetros. Las grandes áreas de color claro corresponden a domos o cúpulas graníticas; son restos de masas de rocas fundidas a consecuencia del calor originado por la radiactividad natural de algunos elementos como el uranio. Las formaciones rocosas alcanzaron la superficie en el Precámbrico (hace más de 600 millones de años). La meteorización posterior del granito aflorante liberó el uranio que se incorporó al ambiente circundante de rocas sedimentarias. Cuando estas eran permeables al agua y concurrían además condiciones favorables, se producían concentraciones locales que podían convertirse en yacimientos. El volcanismo es otro mecanismo de aporte de uranio.

dad de Ciudad de El Cabo, observó que la abundancia de los elementos-traza en los granitos seguía por lo general una curva en forma de campana cuando se representaba el número de muestras frente al logaritmo de la concentración del elemento-traza [*véase la ilustración inferior de la página 38*]. En una escala logarítmica, la curva se ajustaba tan perfectamente a una curva normal, o de Gauss, que condujo a Ahrens a proponer que tal ley de distribución log-normal era una ley fundamental de la geoquímica.

Aunque dudamos que la log-normalidad sea una ley, la hemos tomado sin embargo como un punto de partida para nuestra investigación. Nuestra hipótesis ha consistido en preguntarnos si la distribución del uranio en la corteza terrestre se podría asimilar razonablemente a una curva log-normal de forma de campana. Además, nos hemos planteado si en los yacimientos de uranio que de hecho se explotan aparece la parte de la "cola" de ley alta, o de pendiente ascendente, de esa distribución. Esta hipótesis retiene lo que creemos son los dos principios básicos más importantes de la obra de Hubbert, a saber: el reconocimiento del marco geológico que contiene los yacimientos de interés y el estudio de los informes históricos de la industria del descubrimiento de esos yacimientos. Nuestros descubrimientos, recientemente publicados en un informe voluminoso, del grosor de un libro, que se preparó para el Departamento de Energía de los Estados Unidos, sugieren que el modelo de distribución de uranio en la corteza terrestre y el modelo de la historia de su minería se pueden juntar

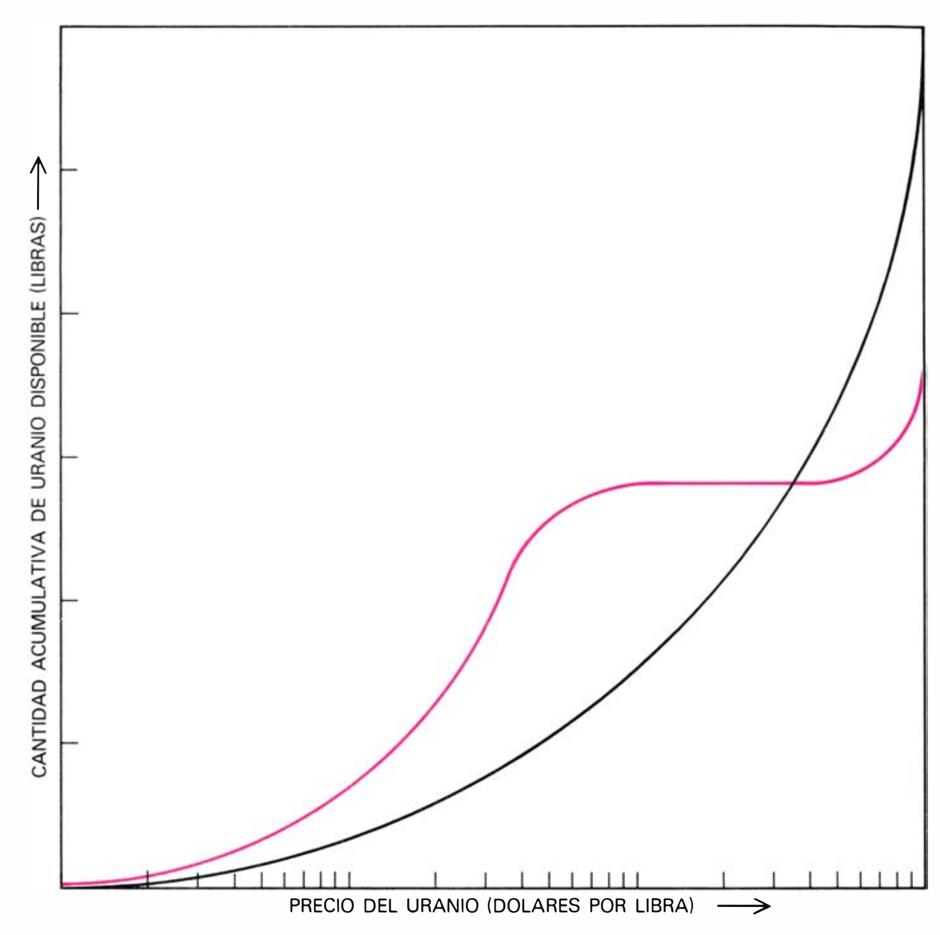


EL URANIO DEL GRANITO se encuentra concentrado en unos pocos minerales raros. Su presencia como fase dispersa en los minerales esenciales del granito es mucho menos frecuente. Esta microfotografía, que se ha hecho con luz transmitida y con un aumento de 50 diámetros, corresponde a una lámina delgada (0,03 milímetros) de granito. Los colores son consecuencia de las interferencias de las ondas de luz transmitida que, al pasar por los materiales cristalinos de la muestra y por el monocristal de cuarzo de la platina, se han separado en dos componentes de distinta longitud de onda. La mayoría del uranio del campo de observación está en un cristal de circón (Zr SiO_3). Es el grano pequeño y de color azul oscuro del centro.

en una figura coherente. En resumen, sacamos la conclusión de que las dos series de datos se pueden representar por una única curva log-normal, cuya pendiente uniformemente ascendente indica un incremento de unas 300 veces en la cantidad de uranio recuperable por cada reducción de diez veces en la ley del mineral. Esta conclusión tiene implicaciones importantes para la disponibilidad futura de uranio.

No hay ninguna teoría con base geológica que exija que la abundancia de los elementos-traza de la corteza terrestre, tales como el uranio, deban seguir una distribución log-normal. Nuestro hilo conductor ha sido la consideración de que los yacimientos ricos en uranio (y en otros elementos-traza) se forman solamente donde se reúnen varias circunstancias poco probables. Tiene que haber: una fuente del elemento, suficiente agua para transportarlo, conductos subterráneos adecuados, agentes acomplejantes para mantener el elemento en solución y otros agentes para precipitar los minerales al final. Evidentemente, si falta uno cualquiera de estos ingredientes no se formará ningún yacimiento de uranio. Por eso, la probabilidad de que se forme un yacimiento de mineral en un sitio dado resulta del producto de las probabilidades de que exista cada uno de los elementos esenciales. Siempre que las probabilidades se suman, el teorema central del límite en estadística muestra que la eventual distribución se aproxima a una curva de distribución normal en forma de campana. Puesto que la multiplicación de las probabilidades corresponde a una suma en escala logarítmica no es sorprendente el ver que la curva campaniforme aparece cuando se representan, en escala logarítmica, las abundancias de los elementos.

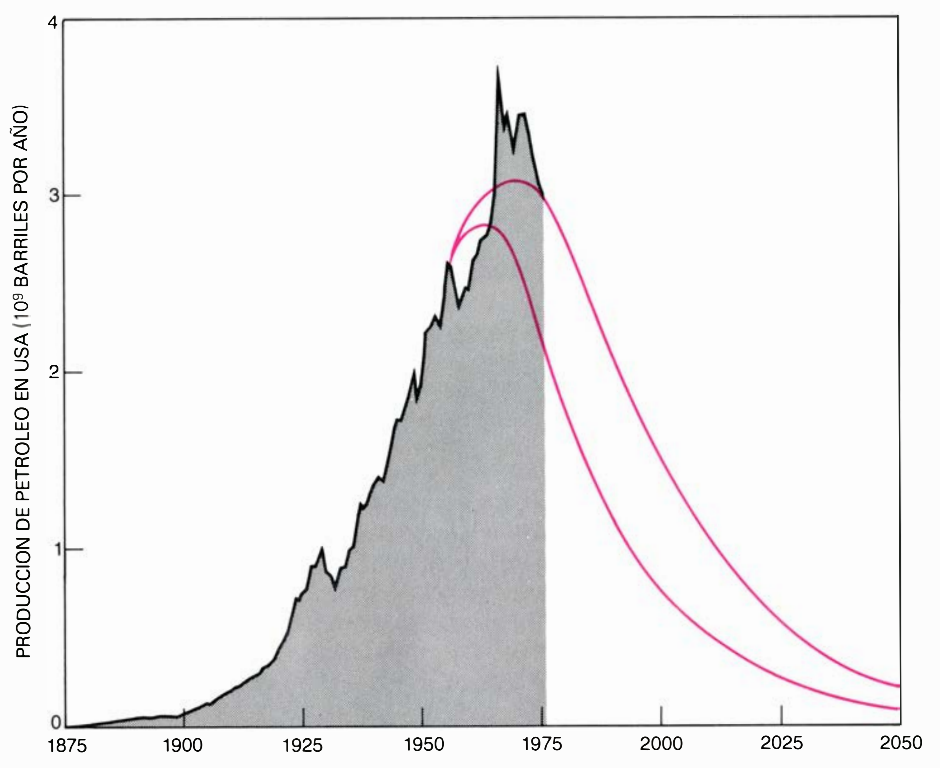
Nos apresuramos a señalar que se trata de un razonamiento de aproximación, pues no existe ninguna base estadística rigurosa para esperar una distribución log-normal. Si la curva que muestra la distribución log-normal fuese una representación exacta de los elementos menores de la corteza terrestre, entonces los yacimientos minerales recuperables serían simplemente la cola de ley alta de la distribución log-normal. Un cálculo rápido muestra que este no es siempre el caso. Por ejemplo, tomando el mejor ajuste de la curva log-normal para el argumento de Ahrens referente a la distribución de plomo en los granitos y extrapolando la curva hasta una concentración de plomo de 20.000 partes por millón (que es la ley límite inferior actual para las menas de plomo recuperables rentablemente), se encuentra que la cantidad total de mena de



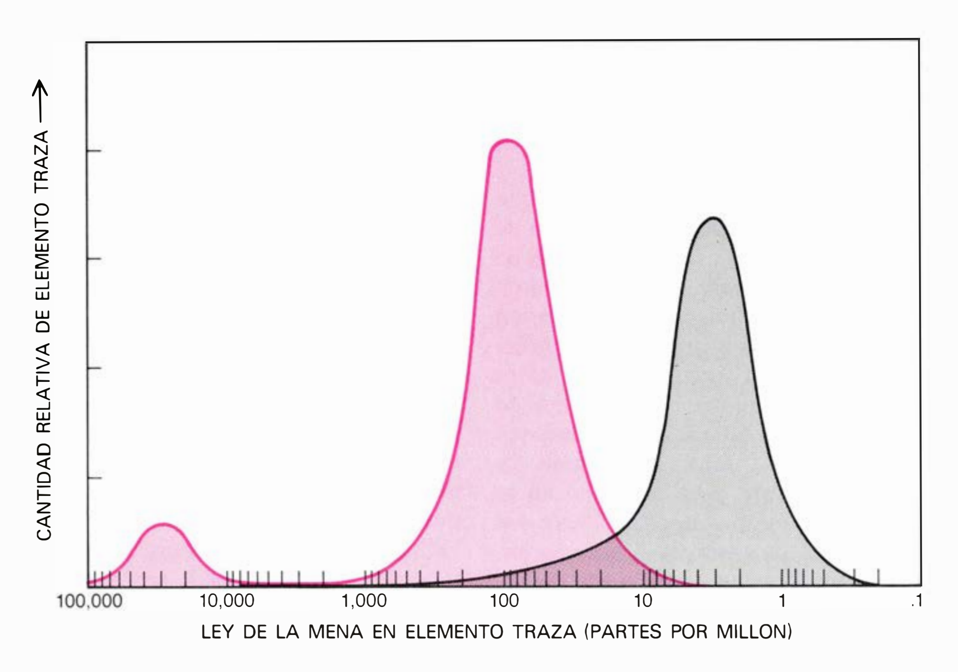
CURVAS HIPOTETICAS DE SUMINISTRO de uranio que relacionan el precio que se prevé para el mismo (*escala horizontal logarítmica*) y la cantidad acumulada de uranio que sería recuperable a ese precio (*escala vertical aritmética*). Una de las suposiciones (*curva negra*) es la de que se siga una trayectoria ascendente uniforme, es decir, que para precios mayores se tendría más uranio. La otra posibilidad supone que un incremento pequeño de precio puede dar origen a una gran disponibilidad de uranio, es decir, que en ese momento se produciría el descubrimiento y explotación de toda una serie de yacimientos de una categoría dada; habría que registrarse, pues, una gran elevación de precios antes que se volviesen a tener otros yacimientos explotables. El resultado sería una curva de suministro que ascendería primero y luego se aplanaría antes de tornar a elevarse (*curva de color*). Los resultados logrados sugieren que la curva correcta es la negra. (Una libra equivale a unos 0,45 kilogramos.)

plomo en el mundo no sería mayor que la de un ejemplar de galena (sulfuro de plomo) de 50 gramos de la colección de minerales de la Smithsonian Institution. Desde luego hay minerales que son tan raros que, de los mismos, sólo conocemos un ejemplar, pero la galena no es uno de ellos; ya se han explotado más de 190 millones de toneladas de plomo que estaba bajo esa forma. Por tanto, la curva de Ahrens para el plomo en el granito se tiene que modificar. Hay dos formas evidentes en que la distribución real de un elemento-traza puede discrepar de una distribución log-normal. O bien la cola de ley alta contiene más material que el que la curva log-normal teórica sugiere, o bien los yacimientos de mineral forman una distribución separada e independiente.

Brian J. Skinner, de la Universidad de Yale, propuso una distribución separada para describir los yacimientos de mineral. Señala que en los yacimientos de menas los metales que se benefician existen como minerales en los que el metal es un componente esencial de la estructura cristalina. En las rocas corrientes las cantidades trazas del mismo metal se pueden acomodar por una sustitución átomo a átomo realizada en el seno de los minerales. Por ejemplo, el plomo existe como galena en la mayoría de las menas de plomo, pero en los minerales que forman las rocas corrientes, tales como los feldespatos, el plomo sustituye en cantidades pequeñas al potasio. Skinner señala que la distribución en el total de la corteza puede ser bimodal, con un pico para los yacimientos de menas y un segundo pico para las rocas corrientes. Las consecuencias para la sociedad serían graves si el pico de ley alta de un elemento-traza como el uranio se agotase y la minería tuviese que cambiar bruscamente a la explotación de rocas corrientes, de ley mucho más baja. Una distribución bimodal del tipo que Skinner propone originaría la clase de discontinuidad representada por la parte plana de la curva coloreada de suminis-



ESTIMACION SATISFACTORIA DE LOS RECURSOS realizada por el geólogo M. King Hubbert en 1957. Trazó una curva de campana para los hallazgos de reservas de petróleo en los Estados Unidos a lo largo de su historia, basándose en las opiniones de los expertos para calcular el área cubierta por la curva. (Las dos curvas de color representan el límite superior e inferior de su estimación.) La marcha reciente de la producción local del petróleo coincide muy bien con las predicciones que hiciera Hubbert. En este caso, el área cubierta por la curva corresponde a la totalidad de los recursos disponibles.



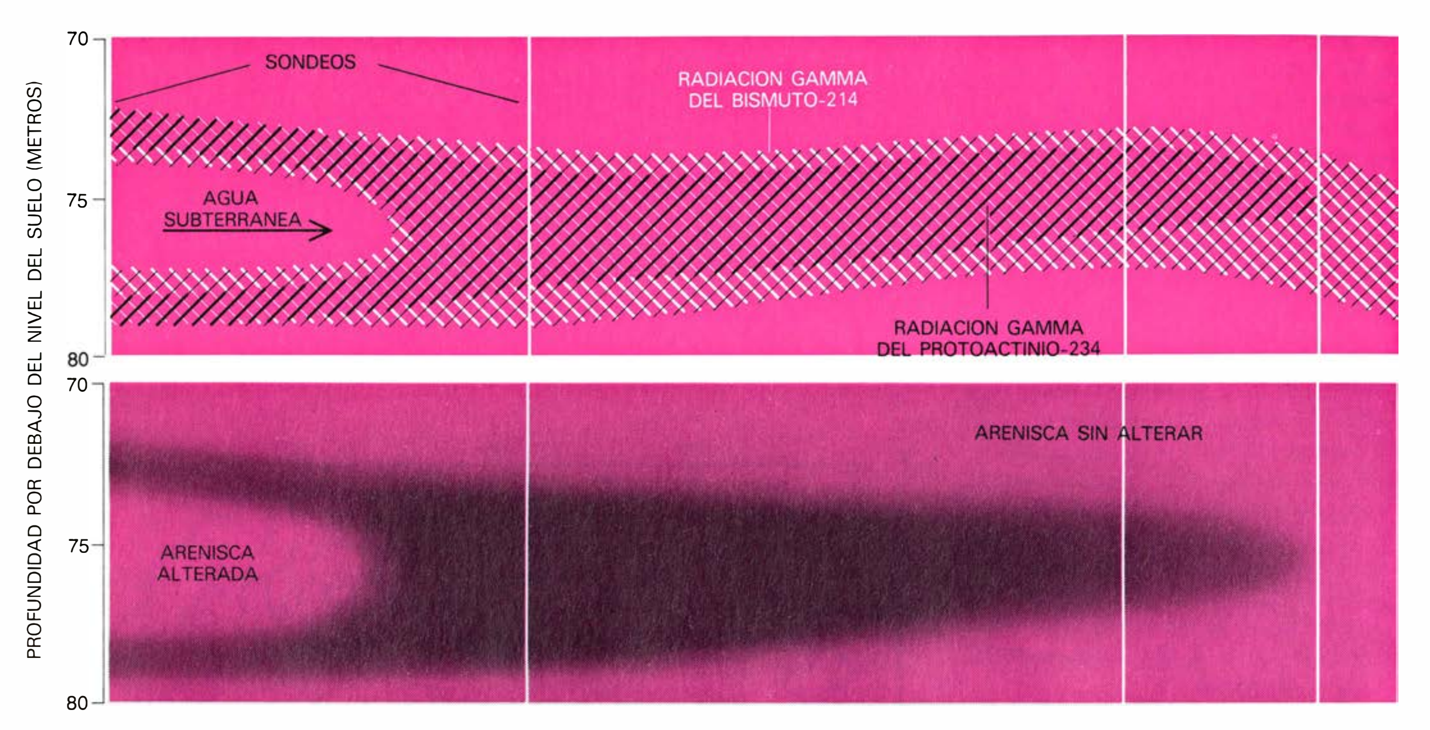
DISTRIBUCION DE LOS ELEMENTOS TRAZA en los granitos. Dicha distribución se adapta generalmente a una curva en forma de campana. En esta figura se han representado las cantidades de elemento-traza de varias muestras frente al logaritmo de sus leyes. La pregunta fundamental que se hicieron los autores al empezar su estudio fue de si sería buena una curva log-normal de este tipo para representar la distribución del uranio en la corteza terrestre. Su interpretación de la realidad parece indicar que podría ser así. La curva negra es el ajuste que lograron para el conjunto de curvas log-normales referentes al uranio. (La curva no es exactamente simétrica: su tramo de leyes altas, o de pendiente ascendente, se ha corregido para incluir algo más de mineral rico que el que correspondería a una curva log-normal teórica.) Brian J. Skinner, de la Universidad de Yale, propuso otro modelo de distribución de los elementos-traza en la corteza; de acuerdo con dicho autor, se tendría una curva bimodal en la que uno de los picos correspondería a los yacimientos de mineral y el otro a las rocas corrientes. El cromo es un ejemplo de distribución bimodal, representado por la curva de color.

tro de uranio que se indica en la figura de la página precedente.

De todos los metales, nos parece que es el cromo el que con más probabilidad muestra el tipo de distribución bimodal que sugiere Skinner. Un solo proceso geológico, la cristalización de grandes masas magmáticas de basalto, ha sido el que ha originado todos los grandes yacimientos de cromo del mundo. La uniformidad de los yacimientos de cromo a lo ancho del mundo puede reflejar la geoquímica relativamente simple de este metal y el predominio de un tipo único de yacimiento.

En contraposición, el uranio muestra un intervalo de comportamiento geoquímico enormemente complejo y una gran variedad de tipos diferentes de yacimientos aprovechables. En la naturaleza el uranio se presenta en dos estados de oxidación, uno con un déficit de cuatro electrones y otro con un déficit de seis, y en una densidad de iones tales como los complejos de uranil-carbonato, $UO_2(CO_3)_2^{2-}$ y $UO_2(CO_3)_3^{4-}$, que contribuyen a la variedad de comportamientos del uranio. El uranio habría de tener una distribución multimodal, en lugar de una distribución bimodal tal como la que se podría esperar para el cromo. Cada tipo químico y cada ambiente geológico tendrá su propia concentración media de uranio característica. Hemos comprobado la hipótesis simple de que los distintos yacimientos de uranio se pueden reunir para formar una distribución log-normal sencilla. Resumiendo, no tenemos ninguna razón teórica fuerte para esperar una curva log-normal; nos hemos limitado a afilar la navaja de Occam y examinar la hipótesis más simple.

Por desgracia no es posible determinar directamente la distribución de los yacimientos de uranio mediante el análisis del uranio en un gran número de muestras tomadas al azar. Aun cuando hubiese un laboratorio que pudiese hacer un análisis de uranio por segundo, se necesitarían 30 años, trabajando contrarreloj, para definir la forma de las colas de la distribución a partir de casos tan raros como las minas aprovechables de uranio. Un esquema analítico de este tipo sería equivalente a prospectar uranio recogiendo rocas a ciegas hasta que se descubriese por casualidad un número importante de minas beneficiables de uranio. Puesto que no es posible medir directamente la distribución, se tiene que analizar el problema por etapas, examinando en primer lugar la distribución de uranio en los yacimientos geoquímicos más importantes y pasando luego a considerar las unidades más pequeñas que mostrasen un contenido de uranio excepcionalmente alto.



LENTO DESPLAZAMIENTO de los depósitos de uranio a través de las areniscas y en el mismo sentido en que fluye el agua, representado en estas dos secciones transversales. Los yacimientos de este tipo, que se denominan yacimientos con frente cilíndrico, tienen zonas de reacción en forma de luna en cuarto creciente. En esta zona se precipita el uranio que hay en el agua y forma uraninita (UO_2). Por medición subterránea de la radiación gamma se puede demostrar el desplazamiento real del yacimiento. Las medidas se hacen con un contador que es capaz de diferenciar la radiación gamma que emite el protoactinio-234 de la que emite el bismuto-214 (*arriba*). El protoactinio es el primer isótopo hijo del uranio-238, mientras que el bismuto-214 es un producto que está mucho más atrás en la serie de desintegración radiactiva. El hecho de que la zona en que predomina la radiación del bismuto (*sombreado blanco*) esté aguas arriba de la zona en que predomina la radiación del protoactinio (*sombreado negro*) se interpreta como prueba de que el yacimiento de uranio (*sombreado inferior gris*) se está desplazando en el sentido de la corriente de las aguas subterráneas que fluyen.

Dos factores dominantes determinan el reparto de uranio en los depósitos geoquímicos más importantes: su gran radio iónico y la producción de calor por radiactividad. En las condiciones probables del manto terrestre el uranio existe ionizado en estado tetravalente, y su radio iónico en los cristales es comparable al radio del ión sodio. Al igual que los iones sodio u otros iones grandes, los iones de uranio entran selectivamente en cualquier fusión parcial que se produzca en el manto. De esta forma los iones uranio emigran con facilidad del manto. (Sucede que los otros dos mayores productores de calor por radiactividad, el torio y el potasio, tienen también radios iónicos grandes y se segregan fraccionadamente del manto por un mecanismo químico. El mundo sería un lugar muy diferente si cualquiera de los iones corrientes más pequeños, el magnesio por ejemplo, fuese radiactivo, pues la fuente de calor habría permanecido entonces más profunda en el manto, en donde habrían dado lugar a corrientes de convección mucho más fuertes y a perturbaciones de una mayor extensión en la corteza.) Los contenidos de uranio extremadamente bajos de los meteoritos de hierro sugieren, por analogía, que el núcleo terrestre también resulta deficitario en uranio.

En la corteza terrestre hay unos pocos nudos cristalográficos en los que el uranio puede encajar adecuadamente. Pero existe una limitación adicional en la distribución del uranio en la corteza. Si el contenido de uranio de un bloque de roca de la corteza genera más calor por unidad de tiempo que el que puede disipar por conducción térmica, el bloque fundirá (de un modo total o parcial) y ascenderá a la superficie como magma fundido. Para ser más precisos diremos que el tiempo que necesita el bloque para alcanzar su punto de fusión es directamente proporcional a su calor específico e inversamente proporcional a la velocidad interna de producción de calor por radiactividad. La distancia a la que se puede difundir el calor está dada por la raíz cuadrada del producto del tiempo por la conductividad térmica. Cualquier masa de roca que esté a la profundidad suficiente y tenga un contenido de uranio lo bastante alto fundirá. En la formación de las montañas, en el choque de dos continentes por ejemplo, los componentes de la corteza recién formada se reagrupan con independencia de su contenido en uranio. Con el tiempo el autocalentamiento puede fundir grandes masas de rocas y formar los granitos postectónicos característicos de las cadenas montañosas de la tierra. El transporte ascendente de los elementos productores de calor, debido a la movilización del granito, es el mecanismo más importante de transferencia del uranio a la parte superior de la corteza continental. Después la erosión expone estos granitos a la meteorización, y con ello el uranio encuentra su camino hacia las rocas sedimentarias.

Hay una segunda forma por la que se incorporan a las rocas sedimentarias grandes cantidades de materiales graníticos uraníferos. Algunas masas fundidas de composición granítica alcanzan la superficie en las erupciones volcánicas. Típicamente las masas fluidas, que son más ricas en agua, forman una espuma de burbujas milimétricas de vapor con paredes de vidrio. Algunas cantidades pequeñas de dicho material se mantienen en este estado constituyendo una roca riolítica ligera, llamada piedra pómez o pumita; sin embargo, en la mayoría de la espuma la expansión continua del vapor hace añicos las paredes de las burbujas, lo que origina una nube de fragmentos finos de vidrio que es lanzada con fuerza hacia la atmósfera. Constituyen las erupciones plinianas, bautizadas así por la descripción que Plinio el Joven hizo de la erupción del Vesubio que sepultó las ciudades de Pompeya y Herculano el año 79 antes de Cristo. En la década de los cincuenta varios geólogos señalaron que la cantidad total de vidrio volcánico, de composición granítica, transferida a la superficie por estas enormes explosiones volcáni-

cas podría tener un volumen similar al de las grandes masas de granito de las cadenas montañosas. Frecuentemente los productos de la erupción se proyectan tan arriba que atraviesan toda la troposfera, y las lluvias de cenizas y de otros fragmentos volcánicos, denominadas genéricamente bombas riolíticas, son extendidas por el viento a lo largo de cientos de kilómetros.

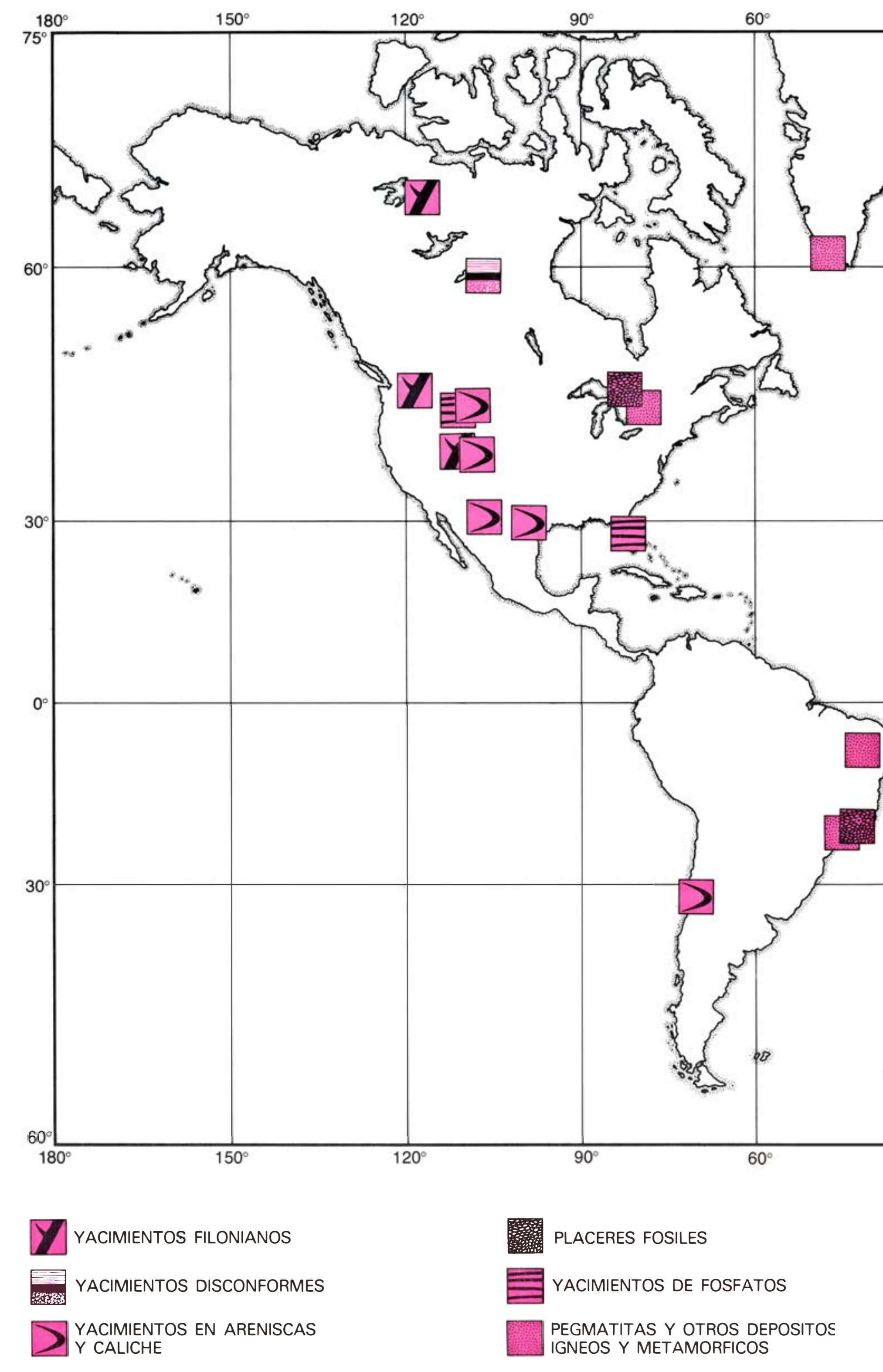
Una vez que la dispersión atmosférica y el transporte subsiguiente, por los ríos, han llevado los fragmentos volcánicos a lugares en los que se acumulan los sedimentos, el vidrio, que es menos estable que una agrupación cristalina de la misma composición, se degrada gradualmente. La liberación de compuestos químicos, en la descomposición de los vidrios volcánicos, explica la extraordinaria fertilidad de los suelos volcánicos. Simultáneamente se libera una pequeña cantidad de uranio.

Las aguas subterráneas pueden transportar, a través de los sedimentos acumulados, el uranio procedente de la meteorización superficial del granito o de la degradación de las cenizas volcánicas; se han encontrado algunos yacimientos de minerales de uranio en los que éste se depositó a partir de las aguas que circulan por el terreno. Hay otros tipos diferentes de yacimientos de uranio, pero de una forma o de otra todos ellos deben su origen a la transferencia ascendente y progresiva del uranio a la parte superior de la corteza continental y también a una o más etapas finales de concentración en la corteza.

El uranio contenido en las formaciones más importantes de la tierra se completa con unidades más pequeñas, que están muy enriquecidas en uranio. De todas esas unidades, las más ricas en uranio son los yacimientos filonianos, fértiles en metal, que se han depositado en las fracturas por la circulación del agua caliente (cuyo calor procedería posiblemente de la radiactividad del uranio de las rocas adyacentes). Los depósitos filonianos tales como los de Joachimsthal en Checoslovaquia, Great Bear Lake en Canadá y Katanga en lo que ahora es Zaire fueron la fuente principal de uranio antes de la segunda guerra mundial. Además de los yacimientos filonianos depositados por el agua caliente, hay algunos yacimientos de uranio de ley alta en pegmatitas; éstas son rocas ígneas de granos anormalmente gruesos, formadas a partir de masas fundidas y ricas en gases disueltos.

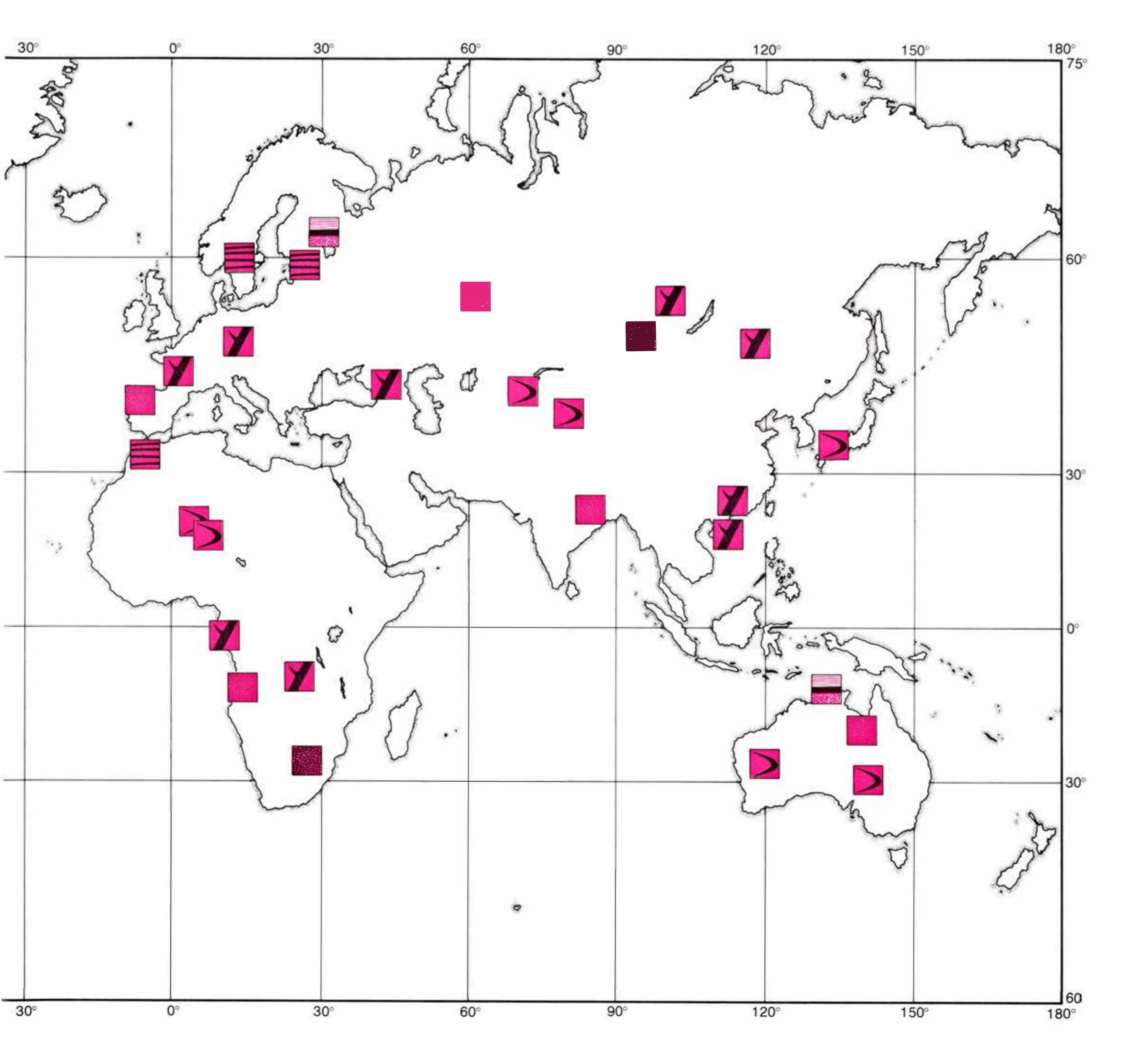
Varios yacimientos de minerales de uranio descubiertos recientemente en Saskatchewan y en el norte de Australia están en la base de capas de areniscas precámbricas que descansan uniformemente sobre granitos y rocas metamórficas precámbricas más antiguas. Posiblemente a causa de su gran antigüedad, estos yacimientos que pertenecen a una categoría de yacimientos de tipo disconforme que se acaba de establecer, se han reagrupado luego en filones o bien han llenado espacios abiertos, y luego se han calentado a temperaturas del orden de unos pocos cientos de grados Celsius. Durante esta última fase de su historia se han borrado la mayoría de las pruebas que habrían demostrado de forma inequívoca su origen. Como consecuencia de ello los geólogos discrepan a la hora de explicar por qué el uranio de estos yacimientos está asociado con las disconformidades.

En orden descendente de concentración en uranio, la siguiente categoría de menas de uranio se compone de yacimientos del tipo arenisca. Estas menas constan típicamente de minerales de

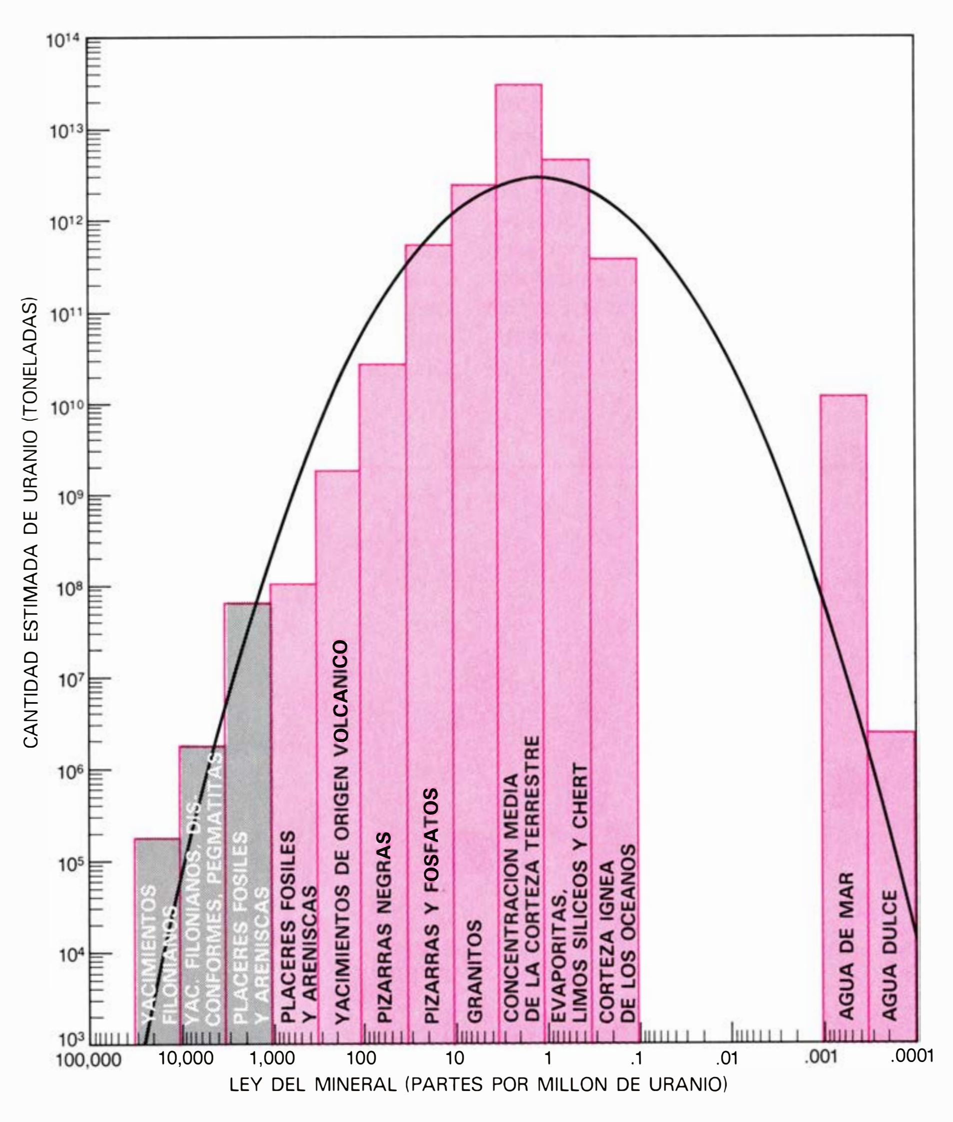


uranio depositados en los intersticios dejados por los granos de arena de las areniscas. El aspecto más extraordinario de tales yacimientos es que muchos de ellos se están moviendo de una forma activa, aunque con lentitud, a través de las areniscas mientras que los mineros de uranio, que se mueven más rápidos, van a su lado y los tratan de pescar. El movimiento del yacimiento de uranio es bastante parecido a la marcha del frente de una columna cromatográfica en el laboratorio. Si en una arenisca que contiene cantidades pequeñas de agentes reductores (tales como materia orgánica fósil, petróleo o sulfuros) se introduce desde agua procedente de la superficie y que tenga oxígeno disuelto, se desarrollará una zona activa de reacción química entre los agentes reductores y el oxígeno del agua. A medida que se siga introduciendo agua rica en oxígeno la zona de reacción se irá desplazando en la dirección del flujo del agua.

La oxidación de sólo al uno por ciento del carbón orgánico de una arenisca requiere todo el oxígeno disuelto en el volumen de agua de 500 poros; por esto, la zona de reacción se mueve más lenta que el agua. Como consecuencia la zona de reacción se convierte en un límite entre el agua rica en oxígeno y el agua de la que se ha eliminado el oxígeno. Aguas arriba de la zona de reacción el uranio se mantiene solubilizado, bajo la forma del complejo de ura-



YACIMIENTOS CONOCIDOS DE URANIO representados en este mapamundi mediante una serie de signos convencionales para los tipos más importantes. La leyenda de la izquierda corresponde en orden descendente de ley a las seis categorías más importantes de menas de uranio. Tanto si se trata de una mena simple como si lo es de una doble (en la que el uranio es sub o coproducto), la explotación de estos yacimientos es rentable hoy o lo será en un futuro próximo. Los yacimientos de placeres fósiles del Witwatersrand, en Sudáfrica, constituyen un ejemplo de mena doble; en ellos se recuperan el oro y el uranio. Los yacimientos de fosfatos también son otro mineral polivalente, cuyo uranio se libera al mismo tiempo que se solubiliza el fosfato como ácido fosfórico. El yacimiento de tipo disconforme que se acaba de descubrir en Jabiluka, al norte de Australia, parece que tiene más uranio que todo el que se produjo en los Estados Unidos entre los años 1948 y 1970. No se han recogido aquí los yacimientos de ley muy baja.



DISTRIBUCION DEL URANIO de la corteza terrestre cuando sólo se tienen en cuenta las cuencas geológicamente más importantes, en una representación doblemente logarítmica. Las barras que corresponden a las distintas categorías de yacimientos (según una ley decreciente) se acoplan a una curva log-normal de abundancia global, que para este caso toma la forma de una parábola cóncava hacia abajo (*curva negra*). La pendiente del tramo ascendente es de 2,5 a 1. Esto equivale a decir que por cada 10 veces que se reduzca la ley de mineral se tendrá un incremento de 300 veces en la cantidad estimada de uranio recuperable. Las tres barras grises indican los yacimientos simples de uranio en explotación.

nil carbonato hexavalente; debido a ello, el uranio que se recoge de la superficie meteorizada del granito, de las cenizas volcánicas alteradas o de otras fuentes diversas se puede desplazar hacia la zona de reacción. En la parte de la zona en la que el oxígeno se ha eliminado de las aguas del terreno, el uranio pasa del estado hexavalente al tetravalente. El uranio tetravalente es insoluble y se deposita por precipitación, generalmente en forma de uraninita (UO_2). Debido a la configuración característica de estos yacimientos de uranio que se mueven activamente, se denominan yacimientos de frente móvil.

Se ha demostrado directamente que los yacimientos de este tipo presentan un movimiento activo en la actualidad. Se llegó a esa conclusión a través de mediciones subterráneas hechas con un espectrómetro de rayos gamma que se introdujo en una serie de agujeros de sondeos. Un grupo de la compañía Princeton Gamma-Tech montó, en una sonda de 5 centímetros de diámetro y 4 metros de longitud, un detector de cristal de germanio, su referencia criogénica de calor y un analizador multicanal. Esta sonda, que el Servicio Geológico de los Estados Unidos, entre otros, ha utilizado para prospectar el uranio, separa la radiación gamma emitida por el protoactinio 234, que es el primer isótopo "hijo" del uranio 238, de los rayos gamma muy abundantes que emite el bismuto 214, un producto posterior en la cadena de desintegración radiactiva.

A partir de los datos así obtenidos podemos representar un corte trasversal que muestra tanto la radiación gamma procedente del protoactinio, que se halla cerca del uranio, como la radiación gamma del bismuto, que, como los diferentes isótopos hijos que se desintegran secuencialmente, ha sido arrastrado aguas abajo por la corriente subterránea [*véase la ilustración de la página 39*]. La nueva técnica es un avance importante; hasta ahora la falta de correspondencia entre la radiactividad del bismuto (predominante en la mayoría de los reconocimientos que miden la radiactividad total) y el uranio constituía un quebradero de cabeza para el explotador, pues las rocas más radiactivas no son forzosamente las más ricas en uranio.

Los yacimientos de uranio del tipo de areniscas de Wyoming, Nuevo México, Utah y Colorado han sido los preponderantes en la producción de uranio de los Estados Unidos; fuera de Norteamérica sólo se conocían unos pocos yacimientos importantes de uranio en areniscas. Varias minas de uranio que se han abierto y que se están abriendo en el Níger tienen, al parecer, el mismo origen que los yacimientos de areniscas del oeste de los Estados Unidos; si bien, hasta la fecha, sólo disponemos de descripciones muy someras de ellas.

La explotación de los yacimientos de uranio de ley más baja que los yacimientos típicos de areniscas no resulta por ahora rentable, a menos que junto con el uranio se extraiga un subproducto o un coproducto. El ejemplo más importante, con mucho, es la producción conjunta de oro y de uranio a partir de los yacimientos del Witwatersrand en Sudáfrica. Durante los primeros 60 años de explotación de los yacimientos auríferos del Witwatersrand los costes de minería y extracción los soportaba ante todo el oro que se recuperaba. La existencia de uranio en las rocas auríferas se conocía de antiguo. Es más, se empleó la radiactividad del uranio como guía de exploración para seleccionar las mejores menas de oro. Después de la segunda guerra mundial empezó la extracción del uranio. Los residuos sólidos de la extracción del oro pasaban a un segundo ciclo para recuperar el uranio. En el proceso de extracción del uranio se liberaba parte del oro que había quedado sin recuperar; esto llevó a invertir el tratamiento; en el "método inverso" el uranio se recupera en primer lugar y el oro en segundo. Los cambios en el panorama económico han conducido a la apertura de algunas minas que no eran rentables para el oro sólo, pero que se hacen viables cuando se explotan con vistas a la producción conjunta de uranio y oro.

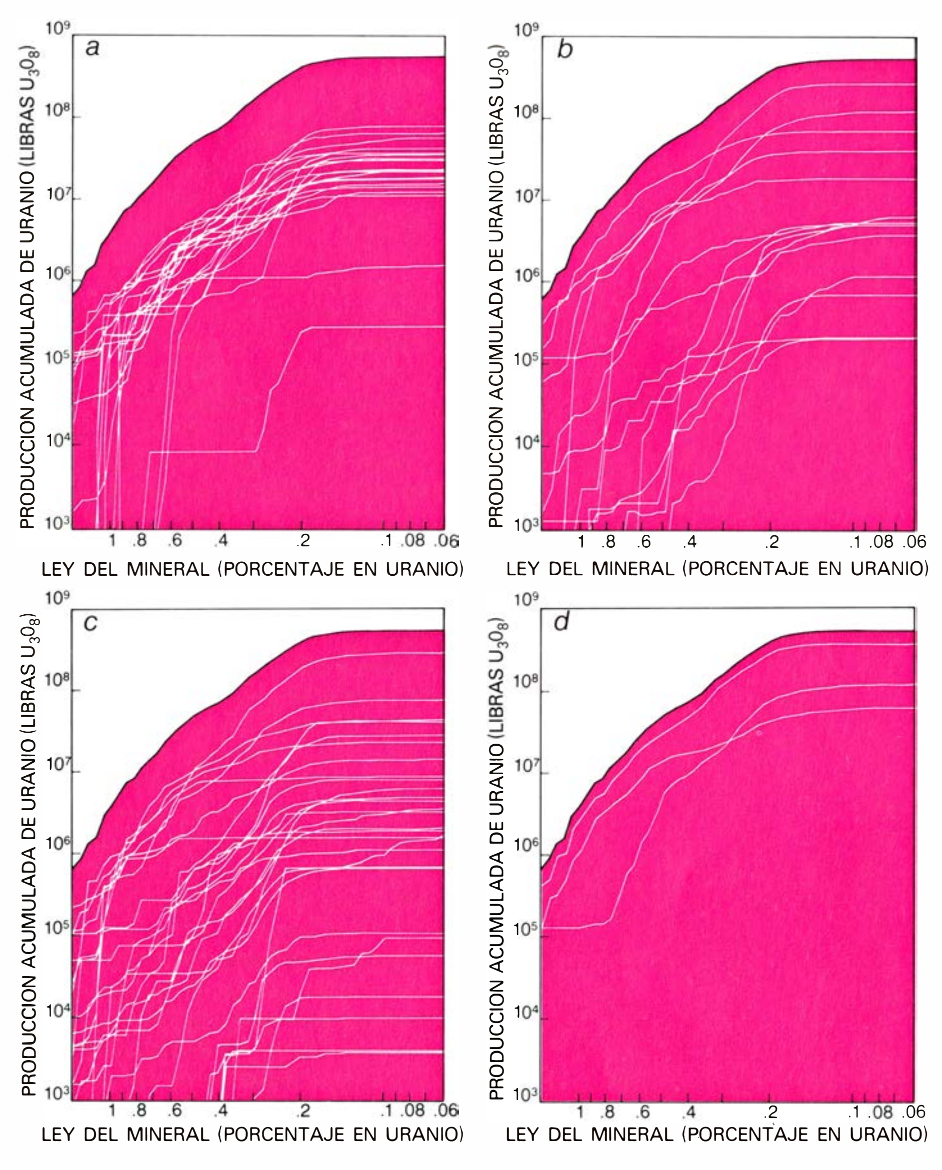
Los yacimientos del Witwatersrand, así como también otros parecidos del Lago Elliot, en Canadá, son placeres fósiles. A pesar de resultarnos familiares

los yacimientos auríferos actuales que se encuentran en placeres fluviales, no sucede lo mismo con el uranio pues la atmósfera oxidante actual de la tierra oxida rápidamente el uranio de los minerales y hace que pase a la forma de complejos de uranio hexavalente que son solubles. Sin embargo, si la atmósfera primitiva de la tierra estuvo en equilibrio con las rocas ígneas de la corteza, las condiciones predominantes debieron haber sido reductoras en lugar de oxidantes. Antes de que la fotosíntesis transformase la atmósfera en su estado oxidante, el mineral llamado uraninita podía desplazarse y rodar por los ríos, ya que eran granos pesados y estables, que se podían concentrar junto con los granos de oro. Ningún yacimiento conocido del tipo de los de Witwatersrand tiene menos de 1800 millones de años. La mayoría de los geólogos aceptan precisamente este descubrimiento como una prueba de que, en aquella época, la atmósfera se estaba volviendo en parte oxidante.

Otra posible asociación que suministrará uranio como subproducto se encuentra en la minería de la roca fosfórica que se destina a fertilizantes. La roca fosfórica típica tiene de 10 a 300 partes por millón de uranio; se trata, obviamente, de cifras mucho más bajas que las leyes mínimas que hoy se están explotando en el beneficio de las menas simples de uranio. Hay que señalar, sin embargo, que en el proceso de transformar la roca fosfórica en fertilizantes fosfatados solubles se libera la mayoría del uranio. Aunque no hubiese otras razones más que las relacionadas con el impacto ambiental, sería conveniente la eliminación o separación del uranio. Una vez que se confirmen las ventajas de recuperar el uranio, se podría valorar la roca fosfórica en base a la suma de los valores del uranio y fosfato contenidos; en consecuencia, podía tomarse en consideración el beneficio de las rocas de uranio-fosfato más valiosas. Iris Y.P. Borg, del Laboratorio Lawrence Livermore, demostró, en 1974, que sólo con el uranio obtenido como subproducto de la explotación de los fosfatos podrían cubrirse todas las necesidades de uranio de las plantas generadoras de energía por fisión de los Estados Unidos.

Existen, además, yacimientos marginales de uranio. Cabe citar, entre otros, los rellenos de las extensas cubetas sedimentarias de Nevada y estados contiguos; en ellas hay mucha ceniza volcánica que suministra uranio al sistema allí existente. Las pizarras negras marinas tienen normalmente más uranio que la media correspondiente a la corteza terrestre; algunas pizarras negras, como la pizarra de Chattanooga, encierran leyes de uranio comprendidas en el intervalo de 10 a 100 partes por millón. Los granitos más ricos en uranio sólo tienen del orden de 10 partes por millón.

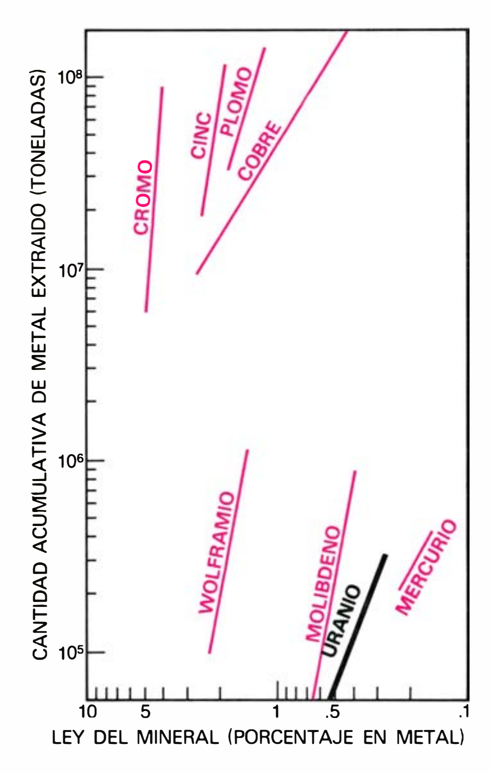
Agrupando todos estos yacimientos pudimos trazar la curva acumulativa de abundancia de uranio [*véase la ilustración de la página anterior*]. Desde luego existe la posibilidad de que hayamos subestimado de forma apreciable la cantidad de uranio en alguna de las categorías. Sin embargo son mínimas las probabilidades de que se hayan subestimado, sistemáticamente, algunas de las categorías más importantes. En apoyo de esto hay que citar que se han estudiado millones de kilómetros de itinerarios de vuelo desde que hace unos 25 años los contadores de radiactividad aérea se convirtieron en un instrumento corriente de prospección. Debido a que la escala vertical de esta representación sui géneris de la abundancia global de uranio es logarítmica, la expresión de la curva normal probable será una parábola abierta hacia abajo. Como muestra la figura, podemos trazar una parábola que se adapte de un modo razonable a los datos. En consecuencia, la hipótesis más simple, la de una distribución log-



CURVAS CALCULADAS POR COMPUTADOR que representan la cantidad de concentrado de uranio (U_3O_8) que se ha recuperado frente a la ley del mineral del que procedía. Se han obtenido a partir de los datos de los informes de un gran número de minas de los Estados Unidos. La curva negra de la parte superior de estos cuatro gráficos representa la historia total de la minería del uranio en los Estados Unidos. Las curvas en blanco se obtuvieron al distribuir los datos disponibles en subconjuntos que tienen en cuenta la profundidad de la mina (*a*), la geología de la roca madre (*b*), la mineralogía del uranio (*c*) y las regiones de los Estados Unidos en que geográficamente estaban las minas (*d*). En todos y cada uno de los casos, las pendientes de estas curvas vienen a ser las mismas: aproximadamente de 2,5 a 1. Esto implica de nuevo que se tendría un incremento de 300 veces en la cantidad de uranio disponible por cada reducción de 10 veces en la ley del mineral. Estas son curvas-resumen representativas de un atlas de 400 páginas que los autores prepararon para el Departamento de Energía. En su opinión, en los Estados Unidos se podrá recuperar gran cantidad de uranio a partir de minerales con leyes comprendidas entre 300 y 800 partes por millón (entre el 0,03 y el 0,08 por ciento de uranio).

normal sencilla, explica la mayoría de los datos, y por ello no vemos ninguna razón que por ahora nos lleve a tomar en consideración hipótesis más complicadas. Tiene un interés muy especial el valor de la pendiente en la parábola para el intervalo de las leyes que se explotan actualmente. La pendiente, en la representación doblemente logarítmica, viene a ser de 2,5 a 1, lo que significa que se tendría un incremento de 300 veces, en la cantidad de uranio recuperable, por cada disminución de 10 veces en la ley del mineral. Esta tendencia parece garantizar un incremento uniforme del suministro del uranio a medida que los precios del uranio justifiquen la explotación de menas de ley más baja. Sin embargo, antes de aceptar esta conclusión, consideremos cuidadosamente la historia de la minería del uranio para ver si se confirma la relación entre suministro de mineral y ley del mismo.

Aparte del uranio, para la mayoría de los restantes metales se ha producido, a lo largo de la historia, un cambio progresivo desde la explotación de los yacimientos de ley alta hacia las menas de ley más baja, extraídas de profundidades mayores y con costes más altos. Por el contrario, algunas de las menas de uranio de ley más baja ya se han estado explotando desde los comienzos de los programas intensivos, inmediatos a la terminación de la segunda guerra mundial. La variación con el tiempo de las leyes de mineral ha sido bastante caótica. Ello se debe a que en los Estados Unidos se establecieron subvenciones a fin de fomentar incluso la minería de menas de ley baja. No obstante, ante la ignorancia de las fechas y las cantidades de uranio recuperado de cada categoría de ley de mineral, hemos reconstruido una historia convencional de ese desarrollo, que transcurre como si se hubiesen explotado en primer lugar las menas más ricas en uranio.



PENDIENTE DE LA CURVA que relaciona la ley en uranio del mineral y la cantidad de mena explotada (*línea negra*). No parece muy diferente de las pendientes de las curvas de explotación de otros metales (*color*). La curva que proponen los autores para el uranio es más empinada que la del cobre, pero menos pronunciada que la del cromo.

En los Estados Unidos el único comprador de uranio, al principio, fue la Comisión de Energía Atómica (AEC). Aunque luego se estableció un mercado libre para el uranio, la AEC y los organismos que le sucedieron han continuado manteniendo el registro de las cantidades extraídas de uranio. El acceso a los informes es difícil, pues las compañías mineras suministraron al Gobierno los datos de que disponían, en el entendimiento de que los mismos se publicarían únicamente según categorías estadísticas y no se identificarían las compañías o minas en particular. A lo largo de nuestra investigación hemos ido redactando programas de computador, distribuyendo los datos por categorías. Luego entregábamos cada programa a un funcionario del Departamento de Energía que introducía en la máquina el programa en cuestión, y luego examinaba la salida, o resultado, antes de dejarnoslo ver. Fueron enormes las dificultades que hubimos de superar para lograr el desarrollo de los programas y clasificar los datos de una forma satisfactoria. En el informe reciente que hemos hecho para el Departamento de Energía hemos puesto un apéndice con un atlas de gráficas que tiene unas 400 páginas y que será una contribución parcial al ahorro de problemas de futuros investigadores.

Aunque no hubiese ningún error humano en la confección de la historia de la minería, hay sin embargo un hecho por el que las leyes registradas no representan exactamente las de la mena in situ. En la minería de los metales en general, y en la del uranio en particular, se tiende a componer macizos de mineral cuya ley es enormemente alta con los de mena de ley más baja que se extraen simultáneamente; de esta forma, se mantiene un producto de características bastante constantes. Algunas minas de uranio tienen escombreras a boca mina, en las que almacenan mena de ley alta y mineral de ley baja; de ellas toman material para mezclarlo con la producción diaria a fin de mantener en el óptimo la alimentación que se envía a la planta de tratamiento. Por eso, en muchas minas, los materiales existentes in situ se componen o promedian, a fin de conseguir una ley única de mineral que se puede mantener sin cambiar durante años. Afortunadamente, la mayoría de las minas de uranio tienen su propia planta de tratamiento diseñada para la ley media de esa mina; en consecuencia no es necesario componer o promediar entre minas distintas. Debido al gran número de minas existentes en los Estados Unidos no creemos que, como consecuencia de la mezcla de menas, aparezcan grandes anomalías en los datos.

La curva negra de la parte superior de cada una de las cuatro figuras de la página precedente representa las leyes de uranio de las menas explotadas en los Estados Unidos frente a las cantidades de uranio recuperadas. La pendiente de la curva de los yacimientos explotados de uranio, que es muy parecida a la pendiente de la curva correspondiente a la corteza terrestre para el mismo intervalo de leyes, 2,5 a 1 (en una escala doblemente logarítmica); esto implica una vez más que el uranio disponible se multiplicaría por 300, si las leyes de las menas se redujesen en un factor de 10. Para probar la solidez de esta conclusión hemos dividido los datos en subconjuntos que tengan en cuenta: la profundidad de la mina, el espesor de la capa mineralizada, la mineralogía respecto a uranio, la geología de la roca encajante y la geografía de la región. En todos y cada uno de los casos fueron esencialmente iguales las pendientes de las curvas que relacionan la ley del mineral con el uranio contenido; había que exceptuar, sin embargo, los subconjuntos que comprendían sólo unas pocas minas. Estamos convencidos de que la concordancia entre los datos de abundancia en la corteza terrestre (cuando se toman en un intervalo amplio de ambientes geológicos) y los datos de la historia en los Estados Unidos establecen una tendencia con la que habrá que contar en la futura exploración y explotación del uranio. La curva sugiere además que se recuperará gran cantidad de uranio a partir de materiales con leyes comprendidas en el intervalo de 300 a 800 partes por millón, y que en este siglo no se explotarán yacimientos tan pobres como los representados por las pizarras de Chattanooga.

La pendiente de la curva que relaciona la ley en uranio del mineral y la cantidad de mena de uranio explotada se puede comparar con las pendientes de curvas similares de otros metales [*véase la ilustración de la izquierda*]. Tal comparación muestra que la curva que pro-

ponemos para el uranio es más empinada que la curva correspondiente al cobre, aunque más aplanada que las curvas correspondientes al cromo y al molibdeno. En consecuencia, no parece que la curva de uranio sea anormal.

Hay un paralelismo sorprendente entre las características del uranio y las de la plata. Ambos metales presentan una abundancia media parecida en la corteza (una parte por millón para el uranio y 0,1 partes por millón para la plata); los dos han tenido precios parecidos (hasta hace poco); para ambos las leyes más pobres de las menas aptas para la minería suponen enriquecimientos de 2000 veces la media de la distribución en la corteza, y los dos se recuperan mediante el uso de un agente acomplejante que disuelve selectivamente el metal a partir de esas disoluciones. La principal diferencia es que, así como la historia del uranio es muy corta, la historia de la explotación de la plata en los Estados Unidos empezó hacia el año 1860 y el último descubrimiento importante de plata fue en 1915. Desde entonces la mayoría de la producción local de plata procede de la minería del cobre, en la que la plata es un subproducto.

Para estimar la cantidad de uranio que es probable que se descubra en cada categoría de ley de mineral, se dispone de otro segundo método. Este se deduce de la extrapolación de Hubbert sobre la cantidad de petróleo que se descubre por metro de sondeo perforado. Michael A. Lieberman, de la Universidad de California en Berkeley, representó en una gráfica la cantidad de uranio que se había encontrado por cada metro de sondeo de exploración perforado frente a la cantidad total de metros perforados de sondeos de exploración. Luego fue más allá y estimó la magnitud total de los recursos por extrapolación, en la curva resultante, del valor que se tendría para una cantidad infinita de metros de perforación en la fase de exploración. Puesto que la curva de Lieberman muestra un descenso aproximadamente exponencial, es muy sencilla la extrapolación matemática de la misma. Sin embargo, hay varias pegas en relación con los detalles del satélite de Lieberman. Los datos de que disponía Lieberman estaban distribuidos en categorías según los costos de explotación en curso, es decir, las cantidades conocidas de uranio que eran beneficiables a 17,6 dólares por kilogramo de U_3O_8 en el concentrado de uranio, a 22 dólares por kilogramo, a 33 dólares por kilogramo y a 66 dólares por kilogramo. Lieberman era consciente de que en unos cuantos años desaparecería más uranio de la categoría de 17,6 dólares por kilogramo por efecto de la inflación que como consecuencia de la explotación. Para salvar esta dificultad, nosotros hemos realizado nuestros propios cálculos referidos tanto a dólares de valor constante, que están corregidos de la incidencia de la inflación, como también a dólares de valor actual.

Un segundo problema relacionado con el análisis de Lieberman era la poca información disponible sobre las menas de ley baja. Lieberman empleó la relación de la cantidad de mena de ley baja (coste alto) a la de ley alta (coste bajo) para evaluar la disponibilidad total de uranio que se podía tener a partir de las menas de ley baja. En nuestro estudio encontramos, sorprendidos, que para un año sólo tres compañías se habían molestado en informar sobre sus reservas de mineral de la categoría de ley más baja (66 dólares por kilogramo). Esta omisión no sólo subestima muy seriamente la cantidad de mena de baja ley disponible, sino que señala también otra estimación por defecto. Tiene en cuenta únicamente la cantidad de mineral de ley baja que hay dentro de los yacimientos de ley alta, o a su alrededor, pero desprecia todos los restantes terrenos que no han producido nunca menas de uranio de ley alta, y en los que se sabe existen yacimientos de ley baja. Por estas dos razones, Lieberman llegó a estimaciones de uranio recuperable mucho menores que los valores deducidos actualmente por el Departamento Norteamericano de Energía.

Nosotros hemos repetido el análisis de Lieberman pero empleando diversas magnitudes para medir el esfuerzo de exploración; no nos hemos limitado a la consideración del número de metros de sondeo perforados, sino que hemos atendido también al número de perforaciones, los gastos totales de exploración y el número de hectáreas de terrenos sobre las que se han concedido permisos para exploración. Cuando estas estimaciones se traducen en dólares constantes y en dólares en curso, se tiene un intervalo completo de estimaciones, que en su mayoría están comprendidas entre la estimación de Lieberman y las del Departamento de Energía. Nuestra impresión es que la dispersión entre las distintas estimaciones da una idea aproximada de la incertidumbre del método en sí, y somos bastante escépticos en que a partir de todas esas respuestas dispersas se obtenga una estimación "correcta".

Las perspectivas de que se producirá un aumento del suministro del uranio a medida de que se rebaje la ley del mineral parecen mucho más optimistas que para la aplicación de esa misma hipótesis al caso del petróleo. Los Estados Unidos podrían abastecer, con uranio producido en el país y sin que el costo de la minería del uranio representase una parte importante del de la energía, a una industria de energía nuclear incluso ampliada; para ello tendrían que emplear la fracción de ley baja de los yacimientos del tipo de areniscas, recuperar el uranio que pudiese resultar como subproducto de la minería de los fosfatos y buscar otras menas comprendidas en el intervalo de leyes de las 600 a las 2000 partes por millón. Cabe la posibilidad de que esto no llegue a cumplirse. Como ya ha sucedido muchas veces en el pasado, es posible que se ofreciesen en el mercado estadounidense concentrados extranjeros, a un precio más bajo que el costo resultante para la producción local. En particular, las leyes de mineral de algunos de los yacimientos de tipo disconforme de Canadá y Australia están en el orden de las 10.000 partes por millón; además, se trata de yacimientos bastante someros que se podrían explotar por métodos de minería a cielo abierto. Se ha visto, a partir de los sondeos con testigo, que en sólo uno de estos yacimientos, el de Jabiluka del norte de Australia, hay más uranio que el que se ha producido en los Estados Unidos en todo el período que va de 1948 a 1970. Si las enormes cantidades de uranio canadiense y australiano saliesen al mercado mundial con precios un poco por encima de los costos de minería y tratamiento, los Estados Unidos se enfrentarían con el dilema de o bien primar un programa activo de exploración y producción local, o bien comprar concentrados importados y mantener las menas locales de ley baja como una reserva para el futuro.

No hace muchos años, enemigos de la energía nuclear sostenían que las estimaciones del uranio no descubierto eran tan bajas que impedirían la construcción de nuevos reactores de potencia. Ahora que se tienen estimaciones mucho mayores de las disponibilidades mundiales de uranio, cabe preguntarse cuál será el impacto probable sobre el futuro de la energía producida a partir del uranio. La generación actual de reactores de agua ligera se concibió en una época en que el uranio era barato, menos de 17,6 dólares por kilogramo de uranio en el concentrado; frente a esto habría que tener en cuenta el precio actual de 88 dólares el kilogramo. En el coste de la electricidad, los intereses de la inversión hecha en el reactor integran un componente mucho mayor que la contribución debida al costo del uranio, y en consecuencia los reactores se diseñaron de forma que se redujesen al mínimo los costes de

capital aunque con ello resultase un aprovechamiento poco eficaz del uranio. (Un colega nuestro los llama engullidores de uranio.) El diseño de la nueva generación de reactores tendrá que conseguir una mayor eficacia respecto al uranio y, en consecuencia, prolongará la vida útil de los yacimientos conocidos de este mineral en un factor de unas cinco veces; eso será válido incluso aunque se prohíba el tratamiento de los combustibles nucleares gastados, en un intento de evitar la proliferación de bombas atómicas. Si son correctas las estimaciones de los recursos uraníferos y las previsiones sobre una utilización más eficaz del uranio, se podría retrasar, hasta mucho después del año 2000, la necesidad del reactor regenerador rápido de metales líquidos. Por otra parte podría haber en perspectiva otro reactor regenerador. Una simbiosis entre el reactor de fusión rico en neutrones y el reactor de fisión tradicional, pobre en neutrones, podría ser más eficaz que cualquiera de ellos sólo.

(Mientras tanto, hay todavía otro recurso substancial y sin descorchar de combustible de fisión, que no debe despreciarse. Recientemente se ha estimado que la "minería" selectiva de parte del uranio y plutonio altamente enriquecidos que los Estados Unidos tienen almacenados en forma de bombas atómicas obsoletas podría proporcionar combustible suficiente para satisfacer la demanda de una parte importante de los reactores nucleares de potencia de la nación).

Tan importante como haber hecho, en esta etapa, la estimación de las disponibilidades futuras de uranio, nos es el convencimiento de que los géologos y geoquímicos tendrán que acometer incluso mejores estimaciones en el futuro, no sólo para el uranio sino también para un espectro amplio de otros recursos minerales. En los Estados Unidos, por ejemplo, en algunos metales como el cobre se está pasando rápidamente a menas de ley mucho más baja, mientras que en otros metales del tipo del plomo las leyes del mineral han cambiado muy poco. Para que una sociedad industrializada como la actual pueda desarrollarse sin conmociones del tipo de la que ha supuesto, en los últimos años, el aumento brutal de los precios de los combustibles fósiles, es necesario que los geólogos estimen o evalúen, con anticipación suficiente, las cantidades de cada mineral que podríamos tener disponibles en el futuro; en esta evaluación tiene que llegarse al detalle de distribuir la cantidad total en distintas categorías de leyes o de dificultad de aprovechamiento.

Ciencia y sociedad

Nobel de economía y tradición normativa en la ciencia económica

La concesión conjunta del premio Nobel de economía de 1979 al británico Arthur W. Lewis y al norteamericano Theodore W. Schultz supone una variante en la orientación general seguida por la Academia sueca durante el primer decenio transcurrido desde la inclusión de los economistas en la lista de los galardonados. Al hacer esta comparación no se trata de discutir aquí las motivaciones que haya podido haber en la concesión de los premios, ejercicio superfluo que no haría más que añadir una convención sobre otra. La intención es señalar sucintamente, con la perspectiva de diez años, las tendencias metodológicas de los economistas elegidos para tratar de situarlos dentro de las principales corrientes de pensamiento de la ciencia económica.

Hasta ahora, se han reconocido sobre todo las aportaciones analíticas realizadas dentro de la denominada economía de equilibrio general. Así podría citarse a Paul Samuelson (1970), John R. Hicks (1972), Milton Friedman (1976), en el terreno de la teoría, o a Jan Timbergen (1969), Wassily Leontief (1973) y Tjaling C. Koopmans (1975), en el de la política económica. Con independencia de su campo específico de estudio, todos estos economistas tienen por denominador común un marco metodológico de referencia construido a partir del análisis neoclásico. Dos casos atípicos en esta secuencia serían los de Kenneth Arrow (1972) y Herbert A. Simon (1978), cuyo enfoque científico desborda el campo estricto de la economía. Arrow se ha adentrado en la filosofía política y en la teoría moral para fundamentar su estudio de la elección social, mientras Simon se ha basado en la psicología del comportamiento y en la teoría de la decisión para construir su análisis de los sistemas de organización. Con la nominación de Lewis y Schultz se premia por primera vez la labor de los especialistas en el campo de la economía aplicada y en particular del análisis normativo del proceso de desarrollo económico, entendido como la conjunción de factores económicos, sociales e institucionales.

Este balance arroja así una amplia mayoría de economistas premiados que podrían encuadrarse metodológicamente en la llamada economía positiva. Esta circunstancia no hace sino reflejar fielmente la mayor implantación institucional de este tipo de economista y el claro predominio académico de tal concepción de la ciencia económica.

La dicotomía tradicional entre economía normativa, entroncada con la *Political Economy* clásica, y economía positiva, con raíces en la *Economics* neoclásica, responde a la ambigüedad del estatuto científico de la economía. Lo económico se presenta simultánea o separadamente como ciencia de las relaciones sociales de producción y distribución, como ciencia de las cosas que se intercambian en el mercado, o ciencia del comportamiento racional de los agentes económicos: administración de bienes escasos suceptibles de usos alternativos. Desde una óptica neoclásica la depuración del objeto de la ciencia económica ha consistido en una separación de medios y fines. La racionalidad económica quedaría así limitada a la pura administración de los medios. Ello ha conducido a una creciente formalización de la teoría económica, con el propósito de eliminar los juicios de valor de sus postulados. El resultado ha sido contradictorio, pues si bien se ha alcanzado una mayor consistencia lógica del razonamiento económico y un considerable refinamiento técnico, la contrapartida ha sido un progresivo alejamiento de la realidad que ha sacado a la economía fuera de la historia. Podría decirse, parafraseando a Schumpeter, que gran parte de la teoría económica contemporánea, a través de este elevado proceso de abstracción, se ha quedado "demasiado audazmente colgada del aire".

El cambio de nombre de la disciplina ha podido contribuir a alterar la imagen pública del economista, pero la pretensión de neutralidad normativa de la economía parece lejos de haberse alcanzado. En la realidad no existe discontinuidad entre medios y fines económicos y éstos aparecen ligados de acuerdo con una pluralidad de normas, expresión de la escala de valores sociales. El objetivo neoclásico supone sustituir esta pluralidad de normas por una norma subyacente universal, el principio de economicidad, convertida en una lógica del cálculo, ajena a la diversidad de valores sociales. La rama más prometedora salida del árbol neoclásico, la economía del bienestar, presenta importantes limitaciones para el análisis de la elección social, al excluir pronunciarse sobre cuestiones claves como la distribución de la renta y las comparaciones interpersonales de bienestar.

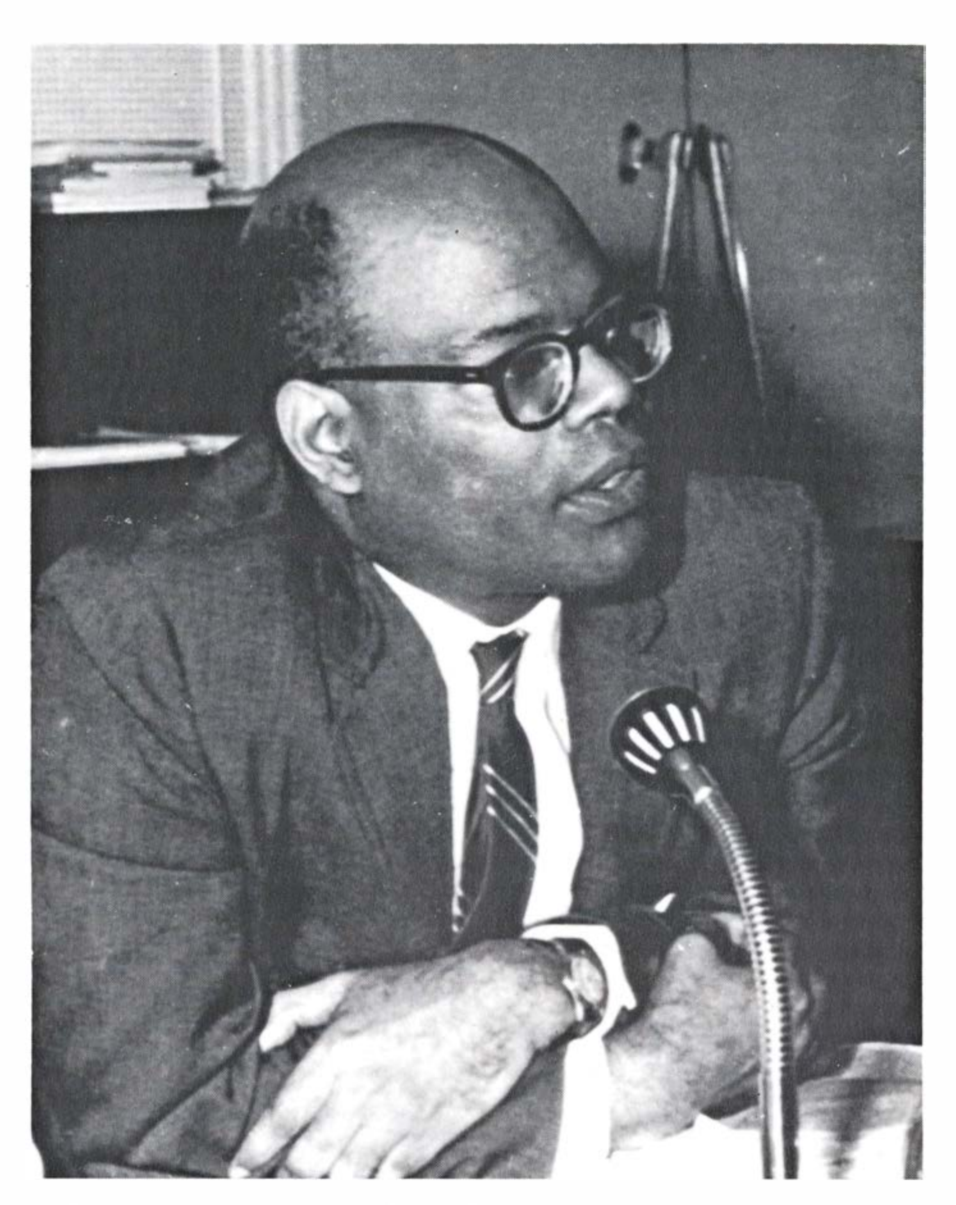
Tras un predominio prácticamente incontestado del pensamiento positivo en la ciencia económica, se asiste en los últimos años a un reflujo metodológico que trata de resituar el análisis económico en una perspectiva normativa. Es lo que se ha llamado una vuelta a los clásicos, entendida en términos estrictos como la recuperación de una manera de pensar la economía de forma global que integre los procesos de producción y distribución. A este giro ha contribuido tanto la creciente insatisfacción por el reduccionismo teórico neoclásico como las presiones de la realidad en forma de profunda crisis económica. La crisis, de características radicalmente nuevas y que alcanza ya dimensión mundial, presenta múltiples e importantes cuestiones normativas, es decir, de elección social, difícilmente ignorables, cuyos fines últimos encuentran su expresión más acusada en la reclamación de los países subdesarrollados de un nuevo orden económico internacional.

El premio a Lewis y Schultz parece enlazar con tal reorientación científica y hacerse eco de la crisis por la que atraviesa la economía mundial. En este sentido puede interpretarse como una especie de reconocimiento oficial de la virtualidad de la tradición normativa en la ciencia económica, no obstante su menor desarrollo teórico en el pasado reciente. La declaración de la Academia sueca así parece indicarlo. Ambos economistas han sido galardonados "por su dedicación común a los problemas del crecimiento en los países en vías de desarrollo", entre los que destacan "las condiciones de eficiencia en la utilización de recursos productivos tanto materiales como humanos, en especial en relación con la agricultura y la educación". La nota señala el carácter complementario de su obra y llama la atención sobre su "amplia labor de estudio de la necesidad y de la pobreza en el mundo, así como su participación en programas económicos orientados a la búsqueda de soluciones al subdesarrollo".

Arthur W. Lewis (Santa Lucía, West Indies, 1915), nacionalizado británico, ha sido profesor de la London School of Economics (1936-1948), Universidad de Manchester (1949-1958), Universidad de West Indies (1959-1962) y de la Universidad de Princeton, New Jersey (USA), donde continúa en la actualidad. Desde 1950 ha simultaneado su trabajo académico con tareas de asesor en las Naciones Unidas y en diversos países subdesarrollados, particularmente del

Africa negra. Durante el período 1970-1973 fue el primer presidente del Banco de Desarrollo del Caribe.

Theodore W. Schultz (Arlington, Dakota del Sur, USA, 1902) ha enseñado desde 1943 en la Universidad de Chicago, en la cual fue director del Department of Economics (1946-1961) y en 1952 fue nombrado Charles H. Hutchinson Distinguished Service Professor; ha sido presidente de la American Economic Association (1960). Ha dirigido o participado en proyectos de investigación y desarrollo sobre países latinoamericanos y asiáticos en colaboración con gobiernos y organismos internacionales.



Arthur W. Lewis ha sido galardonado con el premio Nobel de Economía correspondiente a 1979 (**arriba**). *Lo compartió con Theodore W. Schultz*

Lewis es uno de esos pocos economistas que se ha preocupado de escribir no sólo para sus colegas académicos. Se trata de un excelente divulgador que ha sabido acercar la economía al lector no especializado sin detrimento de la calidad científica. Mediante el contraste de la experiencia histórica y los postulados de la teoría, ha acertado a introducir una serie de temas centrales de la realidad económica contemporánea. Así, la planificación económica (*The Principles of Economic Planing*, 1949); las condiciones del desarrollo económico (*The Theory of Economic Growth*, 1955); los problemas de organización económica e institucional de los países ex-coloniales tras su independencia, expuestos en *Development Planning* (1968), *Politics in West Africa* (1965) y *Some Aspects of Economic Development* (1969); Lewis se ha ocupado también del comercio internacional y las relaciones entre países desarrollados y subdesarrollados, cuya interpretación básica aparece incorporada con nuevas reflexiones en sus dos últimos libros: *Growth and Fluctuation 1870-1913*, (1978), período que está en el origen de la expansión mundial del capitalismo, y *The evolution of the International Economic Order* (1978), síntesis histórica del proceso de división internacional del trabajo que ha conducido a la desigualdad económica mundial actual.

Schultz une, a su formación teórica, un conocimiento práctico del funcionamiento de la economía agraria poco común entre la mayoría de economistas, educados, en general, en un medio urbano. Su extensa bibliografía presenta una cierta concentración temática: las interrelaciones entre agricultura y desarrollo económico y la economía de los recursos humanos.

El núcleo analítico de sus estudios sobre economía agraria se encuentra sistematizado en su obra más difundida, *The Economic Organization of Agriculture* (1953), que ha tenido un desarrollo posterior en numerosos artículos y conferencias la mayoría recogidos en sus libros: *Transforming Traditional Agriculture* (1964), *Economic Crises in World Agriculture* (1965), *Economic Growth and Agriculture* (1968) y *Distortions of Economic Agriculture Incentives* (1978) (Edit.). Schultz ha realizado también importantes contribuciones al análisis del capital humano y en especial a la economía de la educación. Aunque la idea de capital humano se encuentra ya en la literatura clásica y neoclásica, no ha recibido un tratamiento conceptual sistemático hasta un período reciente. El tema recibiría su bautismo oficial con la disertación presidencial del propio Schultz ("Investment in human capital") en la reunión anual de la American Economic Association en 1960. Sus escritos posteriores en este campo, dispersos en numerosos artículos y monografías, están incluidos, en su parte esencial, en los libros: *The Economic value of Education* (1963), *Investment in Education: The Equaty-Efficiency quandary* (1972) (Edit.), *Human Resources* (1972) y *Economics of the Family. Marriage, Children, and Human Capital* (1975) (Edit.). La obra de Lewis y Schultz ha tenido una amplia y continuada difusión mundial y en su mayor parte se encuentra traducida al castellano.

En un momento de desorientación teórica y de aparición de problemas económicos reales inéditos, promovidos por los factores energéticos y ecológicos de la crisis –que han arrojado al baúl de los trastos viejos la mayoría de modelos y políticas de desarrollo– cabe preguntarse cuál puede ser la vigencia de estos dos economistas normativos cuyas propuestas han sido pensadas en gran parte en función de la situación económica de hace tres o cuatro lustros. La respuesta no puede ser unívoca.

Vista desde el Norte industrializado, es decir, los ricos, la obra de Lewis y Schultz –que propugna un desarrollo

equilibrado con participación activa del sector público en la economía nacional y la liberación de los intercambios comerciales internacionales– puede parecer ahora heterodoxa e incluso radical, en contraste con el auge actual del neoliberalismo; nuevo liberalismo limitado a objetivos concretos a pesar de su aparente carácter doctrinario: por un lado, defensa del libre funcionamiento del mercado en el interior (en realidad, desmantelamiento del "welfare state", que sería culpable de los excesos de la política fiscal) y por otro, política proteccionista frente a competencia exterior de las exportaciones procedentes de los países subdesarrollados (apuntalamiento de sectores ineficientes en la agricultura y la industria ante la presión de grupos de interés patronales o sindicales).

Para el Sur, o sea, los pobres, si traducimos el eufemismo geográfico, la novedad de la crisis no hace sino agravar problemas ya viejos (dualismo económico, endeudamiento externo), históricamente enquistados en sus economías. Para estos países, ciertos planteamientos de ambos economistas resultan plenamente válidos, no solo como reflexión crítica sobre la experiencia pasada sino como opciones aún abiertas a la política económica.

Tanto Lewis como Schultz han puesto de manifiesto a lo largo de su obra una serie de tendencias del desarrollo económico contemporáneo que se han revelado cruciales a la hora de caracterizar la crisis presente: hipertrofia del complejo urbano-industrial frente a la marginación del mundo rural y de la agricultura; industrialización intensiva en capital material frente a déficits en capital social y subinversión en recursos humanos; primacía de la maximización de la producción sobre los objetivos de redistribución; concentración de las actividades y de las decisiones económicas y carencia de fórmulas descentralizadoras que fomenten la participación; dualismo a escala nacional e internacional entre industria y agricultura, regiones ricas y regiones pobres. No puede decirse, sin incurrir en una distorsión, que su metodología suponga una ruptura abierta con la teoría neoclásica, pero tanto su campo de estudio como su enfoque analítico muestran claros rasgos diferenciales que permiten relacionarla con otras corrientes del pensamiento económico.

Lewis ha señalado las insuficiencias del análisis económico convencional para el estudio del desarrollo económico: "si se está interesado en el estudio a largo plazo del cambio social o queremos dar cuenta de las diferencias entre grupos sociales y países, tendremos que mirar frecuentemente más allá de la teoría económica contemporánea" (*The Theory of Economic Growth*, p. 13).

Así, ha indicado el escaso valor predictivo de gran parte de la teoría neoclásica y keynesiana, en la situación de los países subdesarrollados. En su lugar, Lewis ha propuesto un modelo interpretativo diferente, mediante una adaptación del enfoque clásico, de "oferta ilimitada de trabajo" (Economic Development with Unlimited Supplies of Labour", *Manchester School of Economic and Social Studies*, mayo, 1954). En síntesis, el modelo parte de una economía prototipo de país subdesarrollado caracterizada por su dualismo (sector subsistencia-sector capitalista), donde el exceso de población respecto a otros factores productivos (tierra y capital) es tal que hay amplios sectores (en su mayoría en la agricultura tradicional) en los que la productividad marginal del trabajo es muy baja e incluso nula. En estas condiciones el salario en el sector moderno o capitalista (no necesariamente privado) viene marcado por el nivel de remuneraciones de fuera del sector, el cual se situará próximo o ligeramente superior al existente en la economía de subsistencia de baja o nula productividad. El resultado es una oferta "ilimitada" de fuerza de trabajo a un coste mínimo mientras a este precio exceda a la demanda. La acumulación de capital en este tipo de economías es función de la tasa a la cual la fuerza de trabajo es reasignada, a un salario más o menos constante, desde el sector subsistencia al sector capitalista, lo que permite incrementar el excedente en este último sector y poner en marcha el mecanismo ahorro-inversión.

El modelo de Lewis, en condiciones de economía abierta, da lugar a implicaciones interesantes para las relaciones económicas internacionales. Se trata entonces de un dualismo a escala mundial donde coexisten un sector subsistencia y un sector capitalista (distribuidos entre diversos países a distintos niveles de desarrollo e integración con el mercado mundial). Uno de los principales hallazgos analíticos derivados del modelo es que la deteriorización tradicional de la relación de precios entre países desarrollados y subdesarrollados no procede de la naturaleza de las mercancías intercambiadas (productos manufacturados por materias primas), sino de las relaciones entre productividad y salarios en las dos zonas, lo que se conoce como relación factorial de intercambio. Estas relaciones de intercambio se basan en el coste de oportunidad de la fuerza de trabajo (bastante más bajo en los países subdesarrollados) y no estrictamente en las relaciones de productividad, a trabajo igual salario igual. Este tipo de relación desfavorable de intercambio, reforzado además por factores institucionales ligados a la dominación colonial, ha sido uno de los elementos explicativos fundamentales, según Lewis, de la evolución histórica de la división internacional del trabajo en los últimos cien años aproximadamente. Aunque acogido críticamente en algunos aspectos, es preciso destacar la influencia seminal del modelo de Lewis, planteado en su versión original hace ya un cuarto de siglo, en la literatura posterior sobre el intercambio desigual de autores como Emmanuel, Arrighi y Samir Amin.

La propuesta de Lewis para lograr una mejora duradera de estas relaciones de intercambio consiste básicamente en una transformación prioritaria de la agricultura tradicional mediante fuertes incrementos de productividad que impulsen un alza de los salarios y, como consecuencia de los precios de exportación, genere un excedente agrario que permita la expansión de los productos manufacturados en el mercado interior y disminuya el déficit exterior al reducir las importaciones de alimentos (paradójicamente gran parte de los países subdesarrollados exportadores de productos agrícolas son importadores netos de alimentos). Sólo así podrá iniciarse un proceso de independencia gradual comercial y financiera respecto a los países desarrollados, en el marco de unas crecientes relaciones económicas y tecnológicas Sur-Sur.

Aunque su enfoque analítico no sea enteramente coincidente, Schultz, del mismo modo que Lewis, recurre a diversos aportes metodológicos con tal de traspasar los límites neoclásicos: "la ciencia económica puede contribuir en gran escala al arte del buen gobierno de la economía nacional; para ello, sin embargo, será necesario redescubrir la economía política; el estudioso ha de ir más allá de la organización de empresas y de las economías domésticas, los mercados de factores y productos merecen también detenida consideración y, finalmente, tampoco debemos excluir los procesos políticos y sociales" (*The Economic Organization of Agriculture*, p. 10). Su metodología es el resultado de una especie de proceso agregativo de diversas corrientes de pensamiento, cuya hipotética área de confluencia estaría delimitada por el institucionalismo, la teoría de la organización y la nueva economía del bienestar. Como consecuencia de este eclecticismo metodológico, su concepción de la organización económica es puramente funcional. Esta puede ser bien "un aparato social para

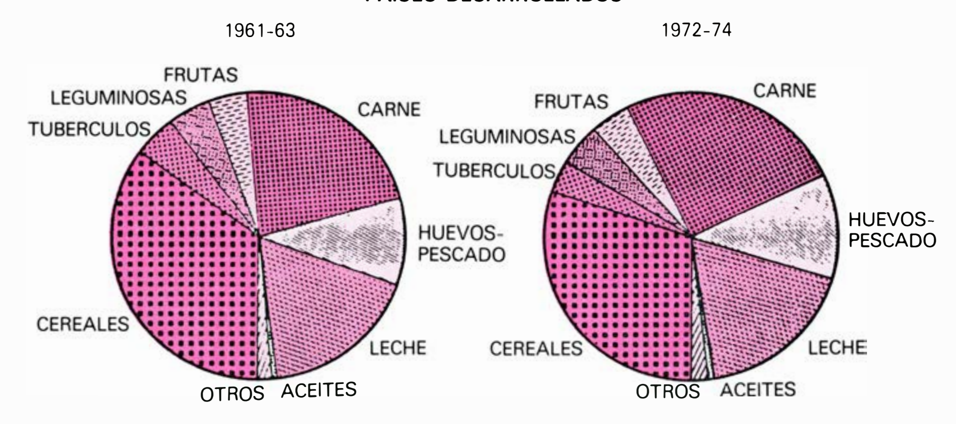
inducir el desarrollo" o bien "un conjunto de instituciones que ajustan la economía a los cambios originados por el desarrollo económico".

Schultz sostiene que la relación de la agricultura con el desarrollo económico no debe limitarse a la contribución pasiva tradicional (trasvase de recursos financieros y de mano de obra, producción de alimentos y mercado de los productos manufacturados), sino que puede realizar un papel potencial nuevo como generadora de una corriente de ingresos que contribuya activamente al crecimiento de la renta nacional. En la situación de la mayoría de países atrasados, el requisito básico para alcanzar este objetivo es una reorientación de las políticas de desarrollo hacia una estrategia equilibrada de la producción entre agricultura e industria, dispersa en la localización de las actividades y descentralizada en la administración de los recursos.

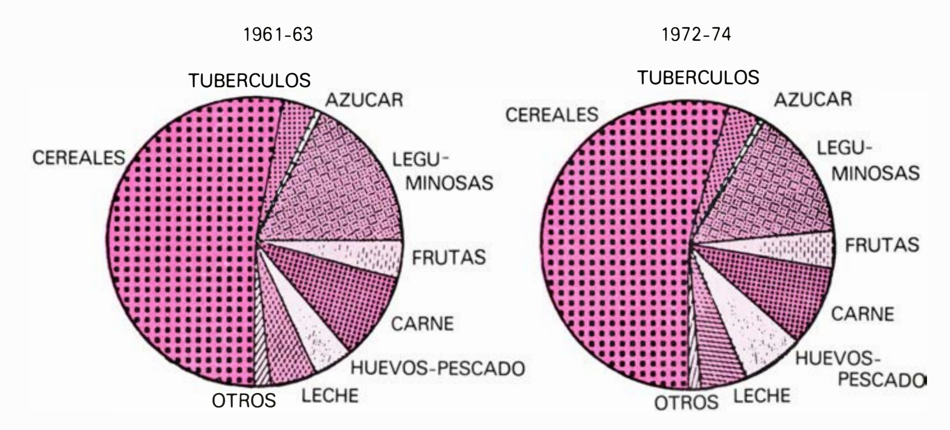
Gran parte de las deficiencias en la asignación de recursos en la agricultura, asociadas con el desarrollo económico, no son consecuencia de la economía agraria *per se,* sino del modelo de desarrollo seguido, orientado hacia una industrialización acelerada intensiva en capital y una concentración de las actividades y de la población en las zonas urbanas. Las consecuencias han sido unas condiciones económicas y sociales desfavorables para la agricultura: depreciación de los productos del campo y carestía de los medios de producción agrarios y de los bienes y servicios que compra la población rural; carencia de medios de financiación apropiados; baja capacitación de la mano de obra; desigualdad de ingresos con otros sectores; grandes déficits de equipamiento social (sanitario, educativo, cultural, etc.). Schultz se ha opuesto firmemente a las críticas formuladas por un cierto fundamentalismo industrialista que hace responsable del atraso y los desequilibrios agrarios a las peculiaridades de la "mentalidad agrarista y campesina" que no respondería suficientemente a los estímulos del mercado. Por el contrario, es precisamente la ausencia o escasez de incentivos económicos y sociales que compensen la introducción de nuevos métodos de producción, lo que obstaculiza, según Schultz, la transformación de la agricultura tradicional.

Su reivindicación de la agricultura no debe interpretarse como un tipo de fundamentalismo agrario de signo opuesto, ni tampoco confundirse con las posiciones de algunos grupos de presión cuya defensa de la agricultura se expresa en una demanda simultánea de protección y de conservación del statu quo. Schultz no ha dudado en apoyar la necesidad, prioritaria en muchos casos, de una reforma agraria allí donde es preciso eliminar la contraposición clásica de tierra sin hombres y hombres sin tierra. Aparte de las razones de justicia social, Schultz ha señalado también la incidencia de este tipo de reformas sobre el modelo demográfico rural mediante la "sustitución de hijos por tierra", según su expresión gráfica. Pero sus preocupaciones fundamentales son para el día siguiente. Es decir, una vez realizados los cambios institucionales, cómo establecer las condiciones precisas para lograr la eficiencia económica de la agricultura en términos de rentabilidad y formas de vida, si no similares, al menos equiparables a la de los otros sectores.

PAISES DESARROLLADOS



PAISES EN VIAS DE DESARROLLO



Porcentaje de los diferentes grupos de alimentos en el suministro de proteínas en los países desarrollados y en vías de desarrollo. (Datos de la FAO.)

El programa propuesto por Schultz consiste en una especie de revolución verde ampliada para cambiar la calidad de los recursos y los métodos productivos en la agricultura, tanto materiales como humanos. Para ello será necesario concentrarse en la formación de dos tipos de capital tradicionalmente menospreciados: la inversión en capital humano (mejora de la capacidad y calificación de la población activa mediante el desarrollo de los servicios sanitarios, educativos y sociales) e inversiones en investigación y difusión de nuevos conocimientos y métodos de producción (mejora en la variedad de semillas, sistemas de cultivo y tecnología apropiada). Dada la actual distorsión del sistema de precios agrarios, la acción pública resultará imprescindible. El objetivo final es lograr un sistema de precios eficientes que permita una asignación de recursos entre los diversos sectores, incluida la agricultura, de acuerdo con las prioridades establecidas por las tasas relativas de rendimiento de la inversión.

En el plano de las relaciones económicas internacionales, Schultz ha sido especialmente crítico del stablishment agrario (privado y público) de los países desarrollados. Así, ha señalado los efectos contraproducentes de la política norteamericana de exportación de excedentes agrícolas que, aparte de paliar situaciones de emergencia de escasez aguda, ha estado guiada a menudo por condicionamientos políticos (como viene a confirmar la reciente decisión de la Administración Carter de suspensión de las exportaciones de cereales a la Unión Soviética, a pesar de la oposición de los agricultores norteamericanos, lo que ha llevado al gobierno a adoptar la decisión, inédita en tiempos de paz, de cerrar

temporalmente las bolsas de contratación de granos).

La ayuda agrícola exterior sería mucho más provechosa si se promoviera la producción nacional de alimentos de los países deficitarios, mediante programas de investigación y extensión agraria, suministro de inputs apropiados y baratos y subvenciones financieras directas a largo plazo a los agricultores y se redujeran las tarifas arancelarias impuestas a las exportaciones agrarias procedentes de los países subdesarrollados.

Schultz se ha ocupado también de algunas de las experiencias de agricultura colectivizada de los países europeos orientales. Su enfoque parte de una evidencia histórica que es a la vez un interrogante: aunque resulta fácil reconocer el pobre rendimiento de esta agricultura —expresado en su repetido recurso a los excedentes agrícolas de los países capitalistas— no está claro de entrada el por qué de su fracaso.

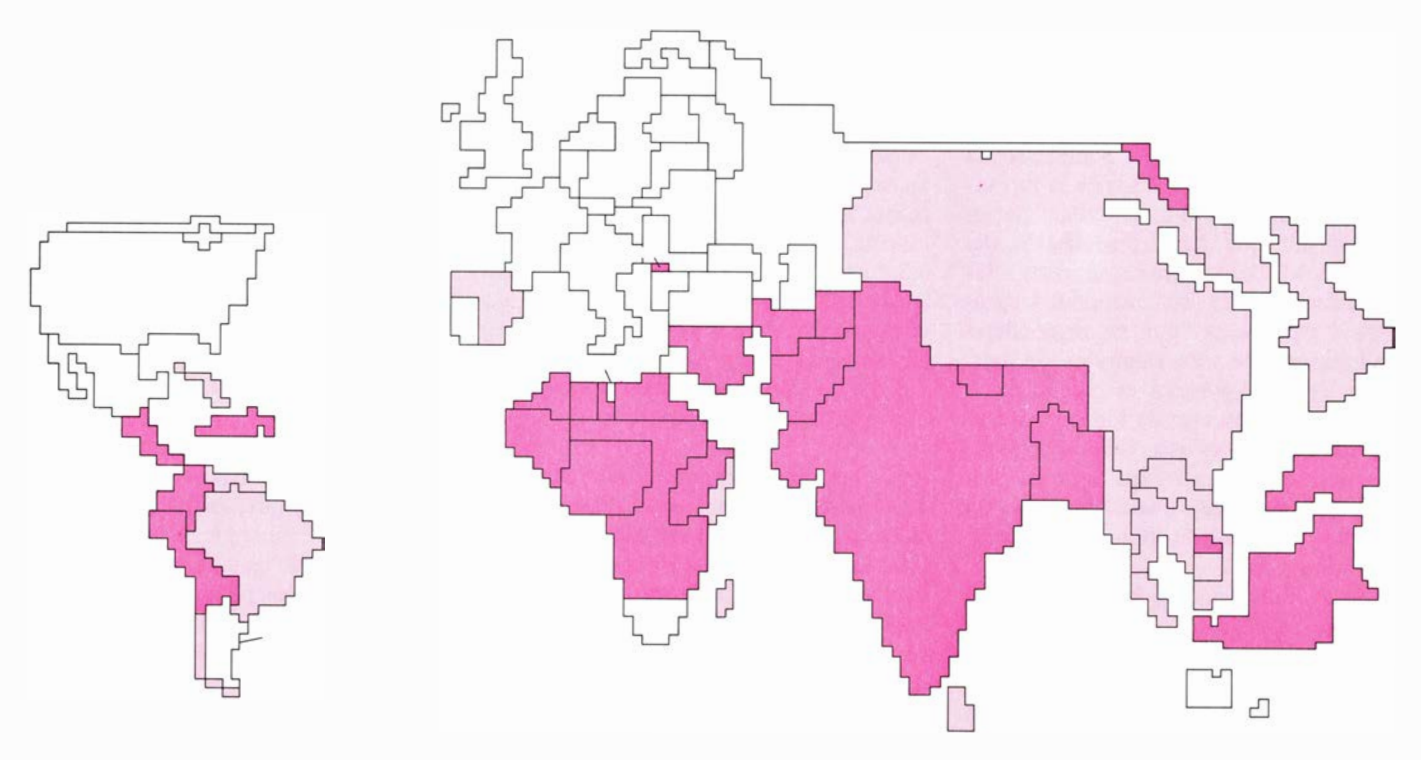
A diferencia de muchos países subdesarrollados aquí no falta ni investigación agronómica, ni tecnología, ni capacitación de la mano de obra, ni infraestructura; lo que falta es una gestión descentralizada por parte de los explotadores directos (se han sustituido las decisiones del propietario absentista por las del planificador absentista) y un sistema de incentivos económicos basado en precios eficientes. Un caso arquetípico de este tipo de ineficiencias estaría representado por la estructura bimodal de agricultura soviética; lo que ha llamado agricultura de "grandes tractores y miles de azadones", polarizada en dos tipos de explotaciones: un grupo de granjas gigantes intensivas en capital, y una miriada de pequeñas parcelas intensivas en trabajo; ninguna de las dos resulta eficiente porque la escala óptima no depende en definitiva del tamaño de la explotación sino de la adecuada combinación de todos los factores de producción.

El principal punto de discordancia entre Schultz y Lewis radica en su diversa interpretación de la llamada teoría de la "productividad marginal cero" en la agricultura tradicional, aunque sus diferencias son más de grado que de fondo. Schultz ha criticado el tratamiento demasiado agregado que se da a la productividad (en la agricultura de subsistencia se cuenta por céntimos no por dólares) y la falta de consideración de la estacionalidad del trabajo agrícola; según Schultz en la agricultura tradicional no existe, en realidad, paro encubierto y la reasignación de mano de obra desde la agricultura a otros sectores debe hacerse mediante un proceso de sustitución de trabajo por otros factores productivos (maquinaria y fertilizantes) si se quiere mantener o aumentar la producción.

Pero estos refinamientos de la teoría no invalidan la proposición central del modelo de desarrollo de Lewis, basada en la existencia de una oferta de fuerza de trabajo, si no "ilimitada", sí considerable a un coste de oportunidad mínimo.

El otro campo temático cubierto por la bibliografía de Schultz es el de la economía de los recursos humanos, entre la que destacan sus estudios sobre el valor económico de la educación concebida en este caso no sólo como consumo de bienes y servicios culturales, sino también como una inversión en capital humano.

A pesar de su connotación economicista y su inicial rechazo moral, el concepto de capital humano resulta sumamente esclarecedor para el análisis del crecimiento económico y el funcionamiento del mercado de trabajo en las condiciones del capitalismo moderno. La concepción del capital como una categoría económica integrada por elementos materiales y humanos (bienes de producción reproducibles y reproducción de fuerza de trabajo calificada) se ajusta mucho más a la realidad que los supuestos de la teoría neoclásica tradicional que trata ambos elementos separadamente y considera el trabajo como un factor de producción homogéneo, simple agregado de la capacidad innata individual. De hecho, la distribución de la capacidad heredada de la población se mantiene prácticamente constante a lo largo del tiempo (dentro de una escala económica significativa) y es más o menos similar entre países ricos y pobres. Es la capacidad adquirida (educación, nivel sanitario, conocimientos técnicos) con un valor económico, es decir, la formación de capital humano, la que altera la distribución de la capacidad natural entre la población y los países.



*Nivel de energía calórica consumida en todo el mundo: inferior al necesario (***color oscuro***), necesario (***color claro***) y superior al necesario (***ausencia de color***)*

La evidencia empírica demuestra que se produce un patrón general histórico de correlación entre la distribución de dicha capacidad adquirida y la distribución del ingreso.

El planteamiento conceptual de Schultz es, en síntesis, el siguiente. Desde un punto de vista público, la inversión en capital humano supone, aparte de una medida de redistribución de la renta, una mejora de la capacidad y de los conocimientos de la población y por tanto contribuye potencialmente a un aumento de la productividad general de la economía; a escala privada (individual o familiar), el rendimiento de la inversión en educación vendría dado por un incremento futuro en el ingreso. Expresada en términos de coste-beneficio, la evaluación de este tipo de inversión implica comparar desde la óptica social/privada la renuncia a la productividad/ingreso presente por la obtención de mayores niveles de productividad/ingresos futuros.

De acuerdo con esta concepción global del capital, la formación de capital humano constituye una de las principales variables explicativas del crecimiento económico, no directamente deducible de la simple acumulación cuantitativa de capital material. El conocido "factor residual", generalmente imputado al cambio tecnológico, vendría así explicado fundamentalmente por un cambio de calidad en la fuerza de trabajo, circunstancia aún más relevante a medida que se avanza en las etapas del crecimiento económico. A diferencia del período histórico de la revolución industrial (intensiva en trabajo de escasa calificación y una técnica relativamente simple), el crecimiento económico contemporáneo se ha basado sobre todo en una mejora simultánea de la cantidad y calidad del capital material y humano, lo cual plantea problemas enteramente nuevos a los países atrasados a la hora de trazar su estrategia de desarrollo. El principal de ellos, el de elección del modelo educativo y tecnológico a seguir.

Tanto Schultz como Lewis han señalado continuadamente a lo largo de su obra el carácter central de esta cuestión. En uno de sus primeros escritos sobre el tema, un informe conjunto preparado para las Naciones Unidas (*Measures for the Economic Development of Underdeveloped Countries*, 1951), establecieron un diagnóstico del subdesarrollo, basado en la subutilización de recursos humanos. Según sus recomendaciones, el principal obstáculo a salvar sería la limitada capacidad de absorción de inversiones directamente productivas, dada la precaria formación de capital humano y los fuertes déficits de equipamiento social existentes; la única vía para alcanzar un crecimiento sostenido sería la de fomentar en forma equilibrada las inversiones en capital material y humano.

Ambos economistas han criticado duramente el síndrome industrialista que aqueja a gran parte de los dirigentes de los países subdesarrollados y la prioridad dada a la obtención de altas tasas de acumulación de capital material frente a la escasa formación de capital humano y la marginación de la agricultura. El resultado de la política de sustitución acelerada de importaciones industriales ha sido doblemente distorsionante. Por una parte, gran número de estos países todavía con mayoría de población rural se han convertido en importadores netos de alimentos, y, por otra, la nueva industria nacional se ha visto sometida a repetidos estrangulamientos, dada su dependencia tecnológica y la limitada expansión del mercado interior, a causa del estancamiento de la agricultura no exportadora.

Esta vía de crecimiento desequilibrado ha inducido una distorsión adicional, de consecuencias aún más importantes a largo plazo: lo que podría llamarse modelo educativo importado, mediante una aplicación mimética, pero de estructura invertida del sistema educativo de los países industrializados. Ello ha dado lugar a una formación de capital humano inadecuado y ha contribuido a acentuar la distribución desigual del ingreso. En el plano educativo: hipertrofia de graduados universitarios, estrangulamientos en la enseñanza media y técnica (jóvenes y adultos) y extrema precariedad de la enseñanza primaria. En el de las remuneraciones el abanico de sueldos y salarios se ha abierto aún más: en un extremo, una oligarquía burocrática pública y privada de alto ingreso según los patrones, incluso superiores, de las grandes metrópolis industriales, y en el otro, grandes masas rurales y suburbanas semianalfabetas con salarios de subsistencia.

A escala individual, el premio obtenido por la inversión en educación es, comparativamente, mucho más alto en los países atrasados que en los desarrollados. Sin embargo, de acuerdo con la experiencia histórica, la mayor contribución de la educación al crecimiento general de la productividad se realiza a medida que la población educada se desplaza hacia abajo en la escala de empleo. Circunstancia que, si bien está avalada por la evidencia empírica, contradice la jerarquía de valores establecidos y choca con la oposición cerrada de las minorías profesionales celosas de conservar sus privilegios económicos y corporativos. Mientras persista tal orientación del desarrollo y continúen existiendo estos obstáculos institucionales será difícil aumentar el rendimiento colectivo de las inversiones en educación.

En la situación común de los países subdesarrollados, una reorientación de la estrategia de desarrollo seguida hasta ahora exigiría, según Lewis y Schultz, una participación activa, aunque no exclusiva, de un sector público renovado (es decir, democratizado), capaz de planificar de una manera descentralizada la formación de capital humano adecuado a las necesidades nacionales, la utilización de tecnología apropiada a las condiciones educativas de la población y la dotación de recursos de cada país.

Desde la perspectiva presente de crisis generalizada de modelos de desarrollo, es decir, de valores, merece la pena hacer referencia a la compartida actitud de estos dos economistas normativos en favor de una integración de los principios de eficiencia y equidad como guía del desarrollo económico. Hace ya casi un cuarto de siglo, Lewis se preguntaba sobre la deseabilidad del desarrollo en estos términos: "la posible ventaja del desarrollo económico no reside en que una mayor riqueza incremente la felicidad, sino en que amplíe la escala de elección humana. No estamos seguros sobre el fin último de la vida, pero en el caso de que este fuera la felicidad, entonces la evolución podría haberse detenido hace tiempo ya que no hay razón para creer que los hombres son más felices que los cerdos o los peces. Lo que distingue al hombre del cerdo no es que sea más feliz sino que tiene un mayor control sobre su medio. A partir de este presupuesto básico el desarrollo económico es plenamente deseable" (*The theory of economic growth*, 1955). Casi simultáneamente Schultz llamaba la atención sobre el sentido elemental del desarrollo cuando éste se contempla desde la pobreza: "se ha dicho que la pluma es más poderosa que la espada; algún día se dirá que las ideas sociales se impusieron a las armas atómicas que nos amenazan; una de estas poderosas ideas es que la pobreza no tiene su origen en un designio divino sino que ha sido creada por los hombres y de ellos depende hacerla desaparecer; el desarrollo económico es precisamente la afirmación de esa idea" (*The economic test in Latin America*, 1956).

Después del paso de una generación algo se ha avanzado por ese camino pero la realidad de la pobreza y la desigualdad mundiales permanece aún con una escandalosa evidencia. Resulta difícil discrepar sobre los hechos aunque su interpretación sea divergente de acuerdo

con los intereses en litigio. Visto desde el Norte, se trata de un problema de exceso de población en el Sur, cuyo crecimiento debería reducirse, para acompasarse a unos recursos limitados. Para el Sur la cuestión reside en un exceso de consumo en el Norte que debería redistribuirse en función de las necesidades de la población y de los recursos mundiales.

En cualquier caso los datos son incuestionables. El 65 por ciento de la población mundial no dispone más que del 15 por ciento del producto bruto total; el grupo de países más industrializados, coincidentes en su mayoría con el de los miembros de la OCDE, poseen una renta per cápita media catorce veces superior a la del resto de países subdesarrollados; se estima que unos 800 millones de seres humanos viven en un estado de pobreza absoluta y que 30 millones de niños menores de cinco años mueren anualmente de hambre; 900 millones de adultos analfabetos ven extremadamente reducidas sus posibilidades de elección social a causa de su pobreza cultural. Mientras tanto, el mundo (el Norte y el Sur) invierte anualmente unos 300.000 millones de dólares en armamento y gastos militares, y la deuda exterior acumulada de los países subdesarrollados no productores de petróleo alcanzará la cifra de 500.000 millones de dólares en 1980; incluso para ese mitológico nuevo "club de ricos" del Tercer Mundo, los países de la OPEP, causante del empobrecimiento del Norte, la situación no ha hecho sino empeorar desde el punto de vista de los intercambios: los ingresos por barril de petróleo de 1975 permitían comprar tres veces más importaciones que los del barril de 1978.

En este mundo profundamente desigual en la distribución de bienes materiales, también en investigación y ciencia existe un Norte rico y un Sur pobre. Esta desigualdad científica y tecnológica ha vuelto a ser puesta de manifiesto recientemente por el co-galardonado premio Nobel de Física de 1979, el profesor de origen pakistaní Abdus Salam, director del Centro Internacional de Física Teórica, con sede en Trieste. El profesor Salam ha señalado la necesidad de acabar con esta división científica entre los que tienen y los que no tienen y que los conocimientos científicos y tecnológicos pasen a ser patrimonio internacional.

Las diferencias actuales no permiten abrigar muchas esperanzas sobre la pronta consecución de tal objetivo. Mientras el Norte gasta anualmente 100 mil millones de dólares en investigación no militar, aproximadamente el 2 por ciento de su renta total, el Sur apenas invierte 2000 millones de dólares en ciencia y tecnología, un porcentaje de su renta muy por debajo de lo que correspondería según los estándares del Norte (cerca de 20.000 millones). Teniendo en cuenta esta realidad, el físico pakistaní ha hecho un patético llamamiento, desde su triple condición de miembro destacado de la comunidad científica internacional, ciudadano de un país pobre y creyente musulmán. En primer lugar, a la comunidad política y científica internacional para que multiplique la promoción y financiación de centros internacionales de investigación para los países subdesarrollados; a continuación ha apelado a la propia responsabilidad de estos países para que detengan la fuga de cerebros y concedan a sus hombres de ciencia la responsabilidad y los medios económicos suficientes para alcanzar al menos los niveles relativos de inversión en investigación y tecnología del Norte; por último, ha recordado a sus "hermanos ricos" de los países islámi-

cos, la posición de vanguardia científica de sus antepasados y les ha instado a destinar una parte sustancial de sus rentas del petróleo a la creación de un fondo de investigación científica y técnica para el mundo pobre. (*Nature*, 13 de diciembre, 1979.)

No está en las manos de los economistas, desde luego, cambiar esta desigualdad económica y científica, pero sí tratar de entenderla y explicarla. Esta es al menos la tarea de algunos de ellos, entre los que se encuentran Lewis y Schultz. El concepto de desigualdad, incluida la económica, es a la vez muy simple y muy complejo; su naturaleza es además histórica. Existe una desigualdad cuya evidencia "objetiva" permite percibirla casi intuitivamente: es una desigualdad que se ve. Aquí basta con el sentido común para discriminar. En otros casos la situación es mucho más compleja y el conflicto entre eficiencia y equidad plantea cuestiones normativas teóricas y prácticas de eleccción social nada triviales. Podría decirse, sin embargo, que una buena parte de las situaciones de desigualdad económica internacional pertenece a la primera categoría de desigualdad simple que está a la vista, no sólo de los economistas. Lo cual no quiere decir que su modificación sea sencilla.

Se ha dicho que las ideas económicas de los políticos en ejercicio suelen ser generalmente patrimonio de algún economista difunto; no debería ser necesario que se cumpliera también en este caso tal aforismo. Algunas de las sensatas recomendaciones de Lewis y Schultz, propuestas desde hace tres o cuatro lustros, resultan ya adecuadas para su aplicación. (Tomás Jiménez.)

El más allá de los griegos

El hombre primitivo sabe perfectamente que un muerto es un objeto inerte, rápidamente corruptible, y hace con él lo mismo que hacemos nosotros: quitarlo de en medio, sea enterrándolo, incinerándolo, exponiéndolo a la acción de las aves carroñeras que dejan el esqueleto mondo y lirondo sin apenas dislocar los huesos, o por cualquier otro procedimiento. Pero lo mismo también que nosotros, se resiste a considerarlo privado de aquello que fue de vivo, aun a sabiendas de que una parte de sus cualidades –sus condiciones personales, sus gustos, su comportamiento– ya no residen en aquel cuerpo. A pesar de ello, deposita al difunto en la tumba, como si la muerte no fuese más que un sueño, en compañía de sus armas, de sus adornos preferidos y de los más escogidos manjares. Sus deudos y sus amigos tratan de demostrarle que persisten los lazos afectivos que los unían con él, mediante todos aquellos regalos de despedida. Los parientes del hombre de Morín, en la montaña de Santander, a pesar de su enorme antigüedad, depositaron junto al cadáver el sabroso costillar de un bóvido, y además un corzo vivo, con las patas atadas como los campesinos solían llevar antaño sus corderos al mercado.

La curiosidad, el afán de conocer, impulsan al superviviente a preguntarse qué será lo que hace vivo a un hombre, esto es, qué será el alma, el espíritu vital. Las posibles respuestas a esta pregunta difieren bastante. Para quienes consideran que la vida se acaba cuando cesa la respiración, el alma es un hálito, un aire particular, el aliento; quienes se fijan más en que el pulso no late y la sangre ha dejado de correr por las venas, identifican el alma con la propia sangre o con el corazón que la bombea, siempre algo cuasi inmaterial y fluido (*spiritus*, en latín). No faltan quienes piensan en un ser

independiente: una mariposa, una serpiente, un pájaro, que se dejan ver cerca de la tumba, o bien en la sombra que acompaña al hombre toda su vida y repite sus movimientos como un *alter ego*... Es admirable el partido que le sacó Chamisso a esta concepción del alma, para su maravillosa historia de Peter Schlemihl, el hombre que vendió su sombra al diablo. Estas y otras respuestas, todas ellas dispares, las formularon y barajaron en sus mentes y en sus escritos los antiguos griegos, hombres curiosos y lógicos en grado sumo.

La concepción más genuinamente griega del Más Allá lo contemplaba como un paraje sombrío, donde las almas vagaban insensibles e inconscientes. Pero al lado de esta creencia generalizada y sustentada sin asomo de preocupación por los poetas más antiguos —"No me alabes a la muerte, Ulises —le dice a éste Aquiles en su visita al Averno—. Más quisiera ser yo el labrador al servicio de un modesto granjero que el Gran Señor en el reino de los muertos" (*Odisea,* canto XI)—, subsistía en algunos la ilusión, heredada del pasado prehelénico, puesta en un país de ultramar donde los difuntos llevaban una existencia similar a la de este mundo, pero más dichosa, sin las penalidades y dolores que aquí experimentan los míseros mortales. Los griegos históricos hicieron de este país de ensueño el Elíseo, un lugar a donde los amados de los dioses, y sólo ellos (no unos hombres cualesquiera), eran trasladados en cuerpo y alma, o lo que es lo mismo, en vida, antes de que la muerte les alcanzase.

Por una serie de vías, apasionadamente exploradas por Erwin Rohde en los capítulos de su obra, *Psique, el culto de las almas y la creencia en la inmortalidad entre los griegos* (Barcelona, Labor, 1973), trataron los helenos de librarse del sino adverso que les imponía la creencia reinante en un mundo de sombras sin consuelo. La tabla de salvación más socorrida fue la de los misterios, sobre todo la de los misterios de Eleusis. Los iniciados en ellos alimentaban la esperanza en una existencia mejor que la de los no iniciados, y no porque estos últimos hubiesen de sufrir castigos en ultratumba, sino por verse privados de la felicidad que las diosas eleusinas otorgaban a los primeros. No faltaron quienes pusiesen reparos a esta creencia: "¿Entonces —preguntaba Diógenes— al ladrón Patecio le va a caber mejor suerte tras la muerte que a Epaminondas, sólo porque aquél ha sido iniciado y éste no?".

Muchos, como los cristianos, no tendrían reparo en que así fuese, con tal de que el ladron se arrepintiese a ultima hora de sus robos y Epaminondas se obstinase en seguir siendo pagano. El criterio tradicional de la Iglesia en este punto está muy bien expresado en el capítulo VIII de la segunda parte del *Quijote*, cuando Sancho pregunta: "Dígame, señor: esos Julios o Agostos y todos esos caballeros hazañosos que ha dicho, que ya son muertos, ¿dónde están agora?".

A lo que respondió don Quijote:

"Los gentiles sin duda están en el infierno; los cristianos, si fueron buenos cristianos, o están en el purgatorio, o en el cielo".

No es fácil de imaginar cómo los iniciados griegos concebían el premio que se les prometía en el Más Allá. Y es que realmente es imposible de concebir una felicidad eterna como un estado de inmovilidad, en una pura contemplación o meditación. Basta con ver las visiones que los artistas cristianos nos dan del Cielo, visiones decepcionantes y aburridas, sobre todo si se comparan con las variadas y pintorescas escenas con que escultores y pintores tan imaginativos como El Bosco son capaces de concebir los tormentos del Infierno. ¡Qué insulsas a su lado las visiones de "La Gloria" de un Ticiano o de un Tintoretto! Tal vez la visión más reconfortante que del Paraíso nos es dado contemplar en el Museo del Prado sea la del frondoso y ameno bosque por donde pasean los bienaventurados en la parte que les corresponde de "La Laguna Estigia" de Patinir. Muy naturales y bien aconsejados anduvieron los escultores compostelanos que labraron en el palacio románico del arzobispo Gelmírez las ménsulas adornadas de aquellos banquetes animados por dulces músicas, aquel saborear de empanadas y de generosos vinos. Es de recordar que Platón sustentaba una creencia muy parecida, pues para él la vida del Más Allá era un simposio en el que los comensales se hallaban en un perpetuo estado de embriaguez.

Si un Homero y un Platón son incapaces de concebir una felicidad consistente en la inactividad es porque la imaginación humana no alcanza a representarse un estado de dicha en que de alguna manera no se repitan las acciones propias de una vida natural ¿Qué ve Ulises cuando visita el mundo de ultratumba? Allí están Agamenón, rodeado de los compañeros que fueron asesinados con él, y Heraklés llevando el arco y las flechas que tanto gustaba de empuñar, y Orión persiguiendo a las fieras que había abatido con su maza por los montes de este mundo. Aun los que aquí se habían hecho célebres por los castigos que los dioses les habían inflingido —un Ticio, un Tántalo, un Sísifo— continuaban padeciendo los mismos tormentos sin variación alguna. El prestigio de Homero había de hacer que estos ejemplos fuesen citados muchas veces y empleados en sus sermones y diatribas por los expositores de las penas del Infierno en otras religiones.

Dentro de esta línea de reiteración, también los iniciados en los misterios griegos contemplaban la vida eterna como una prolongación de su vida natural, con el acto culminante de la celebración de los ritos catárticos. Ahí radicaba la dicha, lo que el no iniciado se hallaba privado de alcanzar, viéndose así reducido a prolongar indefinidamente una existencia vacía y sin sentido.

Antes de llegar a esto, el griego no concebía otra inmortalidad que la del género humano: "La naturaleza mortal procura en lo posible existir siempre y ser inmortal, pero sólo lo consigue mediante la generación, en cuanto ésta deja siempre una cosa nueva en lugar de la vieja". Tal era la doctrina de Platón en el *Banquete* (247 D), inspirada en la imagen del grano de trigo que siglos más tarde había de hacer suya el propio Jesucristo: "Ha llegado la hora de que el Hijo del Hombre se manifieste. Y en verdad, en verdad os digo que si el grano de trigo no muere cayendo a tierra, se quedará solo; pero si muere, dará mucho fruto" (S. Juan XII, 23-25). Tal era el concepto antiguo, consagrado por los misterios eleusinos.

Paulatinamente, sin embargo, y por aplicación gradual del principio de reiteración a que antes nos referíamos, fue abriéndose paso la idea de que los iniciados eran puros antes y después de la muerte, en tanto que los no iniciados conservaban después de ésta su estado de impureza. Esta desigualdad teórica hizo de la inmortalidad un privilegio de los primeros a expensas de los últimos. Así la humanidad como género se vio privado de la inmortalidad, reservada desde ahora a algunos de sus individuos. En los himnos de Prudencio, por poner un ejemplo de un autor cristiano, las espigas verdean, pero no dan fruto; representan solamente la resurrección del individuo.

Para ahondar en aquella diferencia que venía de antiguo, no faltaron quienes propugnaran el justo castigo para los impuros sumergiéndolos en una ciénaga. Con ello los órficos iniciaron una nueva concepción: la de que en el otro mundo tenían tanto espacio los castigos de los réprobos como los premios de los justos. Y al fin quedó muy difuso el saber en qué consistían estos premios, mientras que en cambio los castigos obtuvieron numerosas y prolijas ilustraciones. (A. Blanco Freijeiro.)

Los satélites galileanos de Júpiter

Al avistar los cuatro mayores satélites de Júpiter el año pasado, el vehículo espacial "Voyager" aumentó hasta nueve el número de cuerpos semejantes a la Tierra que pueden compararse estrictamente para comprender su evolución

Laurence A. Soderblom

El día 5 de marzo del año pasado comenzó una de las más espectaculares aventuras científicas de todos los tiempos. En un intervalo de tiempo de alrededor de 30 horas, el vehículo espacial *Voyager 1* pasó cerca del gigantesco planeta Júpiter y envió fotografías en primer plano de tres de los cuatro mayores satélites de Júpiter: Io, Ganimedes y Calixto. Los cuatro satélites se llaman corrientemente galileanos porque fueron observados por primera vez por Galileo en 1610. Más tarde, el 9 de julio, el *Voyager 2* tomó fotografías detalladas del cuarto satélite galileano, Europa, explorando también los hemisferios de Ganimedes y Calixto que habían sido invisibles para su vehículo gemelo. Quizá el descubrimiento más destacado que realizara el *Voyager 1* fue la observación en Io de erupciones volcánicas en progreso. Consiguientemente, el *Voyager 2* dedicó casi 10 horas a la observación de Io haciendo de "vigía de volcanes" en vista de los resultados de las anteriores fotografías.

Ambos vehículos tomaron también fotografías distantes de Amalthea, un cuerpo asteroideo descubierto hace sólo 88 años y que hasta hace poco se creía era el más interior de los satélites de Júpiter. Las imágenes de los Voyager han descubierto ahora, sin embargo, que el planeta tiene otro pequeño satélite cerca del borde exterior de un tenue anillo parecido a uno de los anillos de Saturno. Este catorceavo satélite de Júpiter, cuyo diámetro es sólo de unas docenas de kilómetros y que se encuentra aproximadamente a mitad de camino entre la superficie del planeta y la órbita de Amalthea, fue descubierto por David Jewitt y G. Edward Danielson, del Instituto de Tecnología de California. Provisionalmente se le ha designado 1979J1.

Amalthea y los cuatro satélites galileanos (y quizá también 1979J1) describen órbitas circulares en el plano ecuatorial de Júpiter; constituyen, por tanto, el sistema "normal" de satélites del planeta. Los ocho restantes satélites de Júpiter confirmados son mucho menores y se mueven en órbitas irregulares distribuidas mucho más allá de las de los seis interiores. Del examen detenido de los varios cientos de fotografías de los satélites galileanos y de Amalthea enviadas por los dos Voyager se puede deducir gran cantidad de información acerca de su historia, sus edades relativas y la naturaleza de los procesos geológicos que han determinado su evolución.

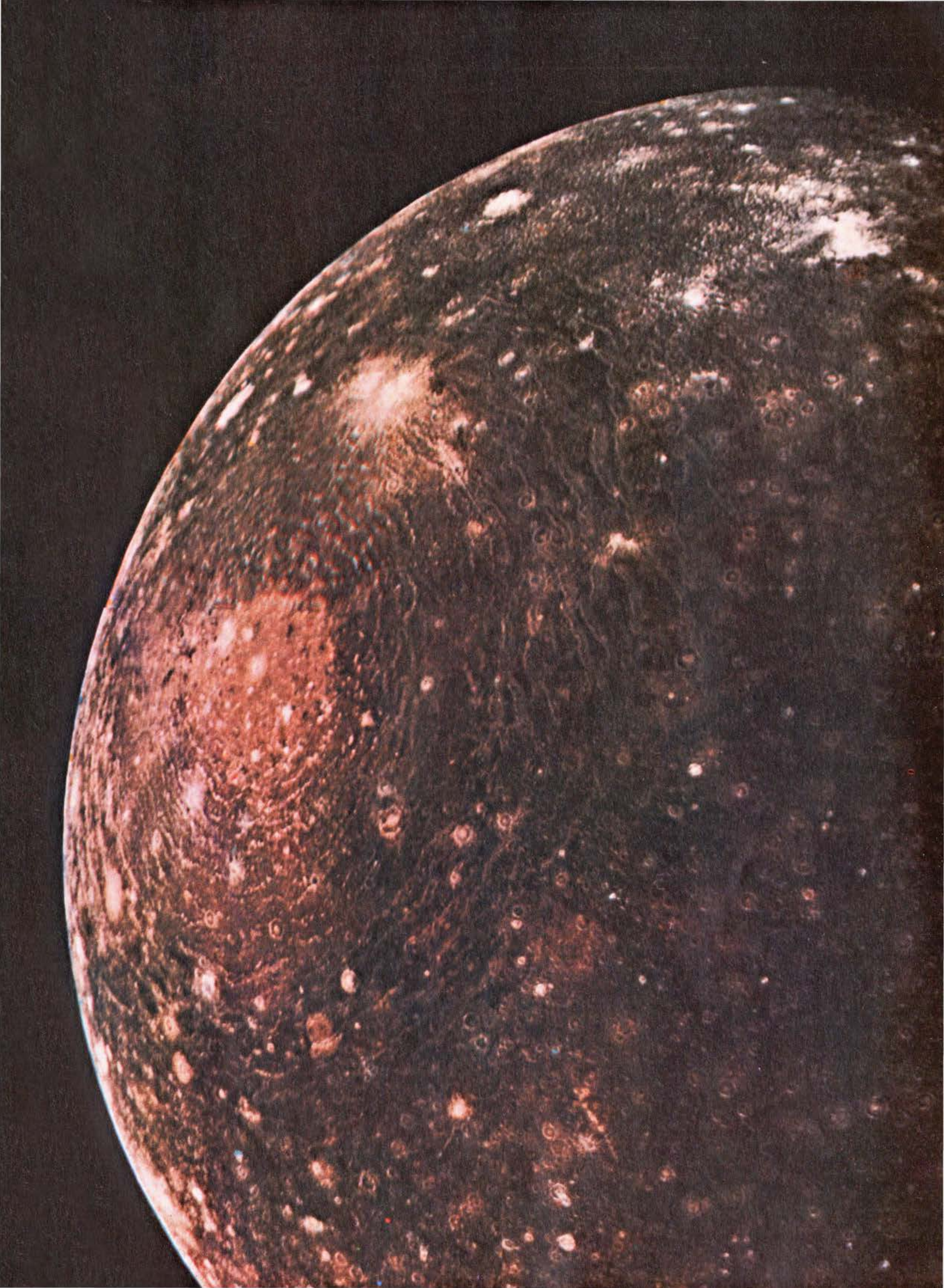
El uso de la palabra satélites para designar a los galileanos subestima su importancia para los estudiosos de las ciencias planetarias. Io, Europa, Ganimedes y Calixto pertenecen a la familia de objetos llamados terrestres, familia que incluye Mercurio, Venus, la Tierra, la Luna y Marte. Los satélites galileanos son semejantes a los cuerpos del Sistema Solar interno tanto en tamaño como en composición, de modo que ambos grupos de objetos deben haber evolucionado mediante procesos comparables y en escalas de tiempo semejantes. Como resultado de las misiones Voyager, se ha duplicado el número de objetos semejantes a la Tierra con los cuales es posible comprobar modelos teóricos de evolución planetaria. Es posible darse cuenta ahora de que los planetas del Sistema Solar interno ocupan sólo una pequeña parte del espectro de características y posibilidades de evolución abiertas a tales objetos.

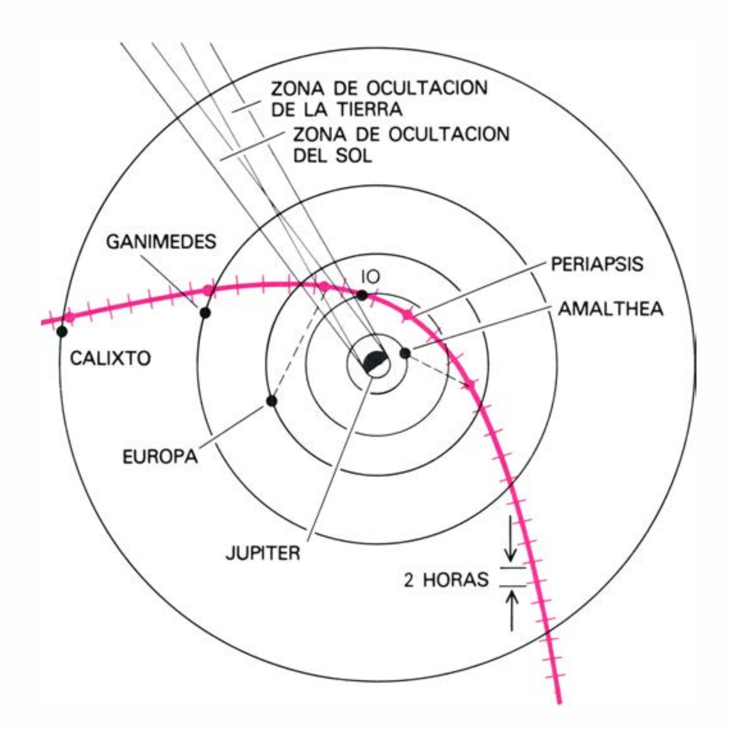
Se sabía, a partir de estudios realizados con telescopios instalados en la Tierra, que los satélites galileanos muestran ciertas características regulares, como son la densidad decreciente y el tamaño creciente con la distancia a Jupiter, características notablemente parecidas a las que presenta el Sistema Solar en conjunto. Desde la Tierra, poco se puede decir de Amalthea excepto que es oscura y roja y que su diámetro mide escasos cientos de kilómetros. Io, el siguiente satélite contando desde Júpiter, tiene, según ya se sabía, aproximadamente la misma densidad (3,5 gramos por centímetro cúbico) y tamaño que la Luna terrestre. A diferencia de los otros tres satélites galileanos, Io no muestra rastros de agua en su espectro infrarrojo de reflexión. Se sospechaba, a causa de su brillante apariencia rojo-anaranjada y el fuerte descenso en su espectro de reflexión hacia el ultravioleta, que poseía una superficie rica en azufre.

Europa, segundo de los grandes satélites galileanos a contar desde Júpiter, tiene también el tamaño de la Luna terrestre, pero es mucho más brillante: refleja casi el 70 por ciento de la luz solar incidente, comparado con el 7 por ciento de la Luna (y el 35 por ciento de la Tierra). Los espectros obtenidos en las longitudes de onda del infrarrojo próximo indican grandes cantidades de hielo en la superficie del satélite. Las estimas de la densidad de Europa (unos tres gramos por centímetro cúbico) sugieren que pudiera tener una costra de hielo y agua líquida de hasta 100 kilómetros de espesor.

Los miembros tercero y cuarto del sistema de satélites galileanos, Ganimedes y Calixto, son bastante semejantes. De ambos se sabía que tenían un diámetro similar al del planeta Mercurio y una

CALIXTO, el más externo de los satélites galileanos, fue fotografiado el 5 de marzo de 1979 a distancias entre 337.000 y 364.000 kilómetros por las cámaras de televisión a bordo del vehículo espacial *Voyager 1*. Los otros tres satélites galileanos, en orden de su distancia a Júpiter, son Io, Europa y Ganimedes. Júpiter tiene en total 14 satélites, incluyendo uno muy pequeño descubierto por el *Voyager 2* y designado provisionalmente 1979J1. El nuevo satélite gira alrededor de Júpiter dentro de la órbita de otro pequeño satélite, Amalthea, que a su vez gira en torno al planeta dentro de las órbitas de los satélites galileanos. Las órbitas de los ocho satélites de Júpiter restantes quedan fuera de las de los satélites galileanos. Calixto es más o menos del tamaño de Mercurio. Da una vuelta alrededor de Júpiter cada 16,69 días a una distancia de 1,8 millones de kilómetros. En este mosaico, la resolución, equivalente a la anchura de un par de líneas de barrido de televisión, es de siete kilómetros. La figura circular cerca del borde izquierdo tiene unos 600 kilómetros de diámetro. Los anillos concéntricos se extienden hacia fuera hasta otros 1500 kilómetros. Cortesía del Jet Propulsion Laboratoy, Cal Tech.





TRAYECTORIAS DE LOS DOS VOYAGER. Ambas se eligieron de manera que su cobertura de los satélites galileanos fuera óptima. El *Voyager 1* (*izquierda*) se acercó hasta 277.000 kilómetros de Júpiter el 5 de marzo. El *Voyager 2* (*derecha*) no llegó tan cerca del planeta, pasando a una distancia de él de 650.000 kilómetros el 9 de julio. Cada satélite está en la posición que ocupaba cuando los vehículos alcanzaron su punto más próximo a él.

densidad de unos dos gramos por centímetro cúbico (las de Mercurio y la Tierra son aproximadamente de 5,5 gramos por centímetro cúbico). Se puede por tanto especular que Ganimedes y Calixto son todavía más ricos en agua que Europa, admitiendo que los tres están formados principalmente por agua y silicatos típicos (óxido de silicio). Ganimedes, sin embargo, refleja el 50 por ciento de la luz solar, comparado con solamente el 20 por ciento de Calixto, y muestra más indicaciones de la existencia de agua o hielo en su superficie.

Las trayectorias de los Voyager

Al moverse a lo largo de sus órbitas, Amalthea y los cuatro satélites galileanos mantienen una cara constantemente dirigida hacia Júpiter, de igual manera que la Luna terrestre mantiene una cara hacia la Tierra. Sus períodos orbitales varían desde 12 horas, para Amalthea, hasta 16,69 días, para Calixto. Las trayectorias de los Voyager aprovecharon esta rotación sincrónica de los satélites galileanos y sus posiciones rápidamente cambiantes para optimizar la cobertura fotográfica de sus superficies. Mientras el *Voyager 1* se acercaba a Io, el vehículo espacial tomó fotografías de la cara exterior del satélite y de los hemisferios intermedios, voló bajo su Polo Sur y continuó para pasar cerca de Ganimedes y Calixto, observando sus hemisferios enfrentados a Júpiter y volando sobre ellos a altas latitudes de modo que se pudieran examinar los respectivos Polos Norte.

El *Voyager 2* encontró a Calixto y Ganimedes antes de alcanzar a Júpiter y por ello pudo fotografiar sus respectivos hemisferios exteriores, opuestos al planeta, y ver la región polar meridional de Ganimedes. Como resultado, se cubrió el 80 por ciento de las superficies de Ganimedes y Calixto con una resolución igual o superior a cinco kilómetros. (La resolución se define como la anchura de dos líneas de barrido de la televisión.) El *Voyager 2* fotografió también alrededor de una cuarta parte de la superficie de Europa con una resolución similar. Como las revoluciones de Amalthea e Io en torno a Júpiter son rápidas comparadas con el tiempo durante el cual los vehículos espaciales estuvieron cerca del planeta, se pudieron fotografiar a todas las longitudes con resolución intermedia

SATELITE	DIAMETRO (KILOMETROS)	DISTANCIA MEDIA A JUPITER (KILOMETROS)	PERIODO ORBITAL (DIAS)	DENSIDAD (GRAMOS POR CENTIMETRO CUBICO)	MASA (LUNA = 1)	MAXIMO ACERCAMIENTO (KILOMETROS)		MAXIMA RESOLUCION (KILOMETROS POR PAR DE LINEAS)	
						VOYAGER 1	*VOYAGER 2*	*VOYAGER 1*	*VOYAGER 2*
AMALTHEA	155 × 270 (±8)	109,900	.49	?	?	420,100	558,270	7.8	11
IO	3,638 (±10)	350,200	1.77	3.53	1.21	18,640	1,127,920	1	21
EUROPA	3,126 (±10)	599,500	3.55	3.03	0.66	732,270	204,030	33	4
GANIMEDES	5,276 (±10)	998,600	7.16	1.93	2.03	112,030	59,530	2	1
CALIXTO	4,848 (±10)	1,808,600	16.69	1.79	1.45	123,950	212,510	2.3	4

CARACTERISTICAS FISICAS Y ORBITALES de Amalthea y los satélites galileanos, tabuladas junto con la mínima distancia de los dos Voyager y la resolución de las mejores imágenes. Cada vehículo espacial llevaba dos cámaras Vidicon con lentes de diferente distancia focal: 200 y 1500 milímetros. Las imágenes en color se producían haciendo exposiciones sucesivas en el anaranjado, verde, azul, violeta y ultravioleta mediante filtros adecuados. Cada imagen consta de 800 líneas de barrido con 800 pixels (unidades de imagen) en cada línea. Los dos Voyager enviaron en total 35.000 fotografías. Todas las publicadas en este artículo se prepararon en el Laboratorio de Propulsión a Chorro del Instituto de Tecnología de California.

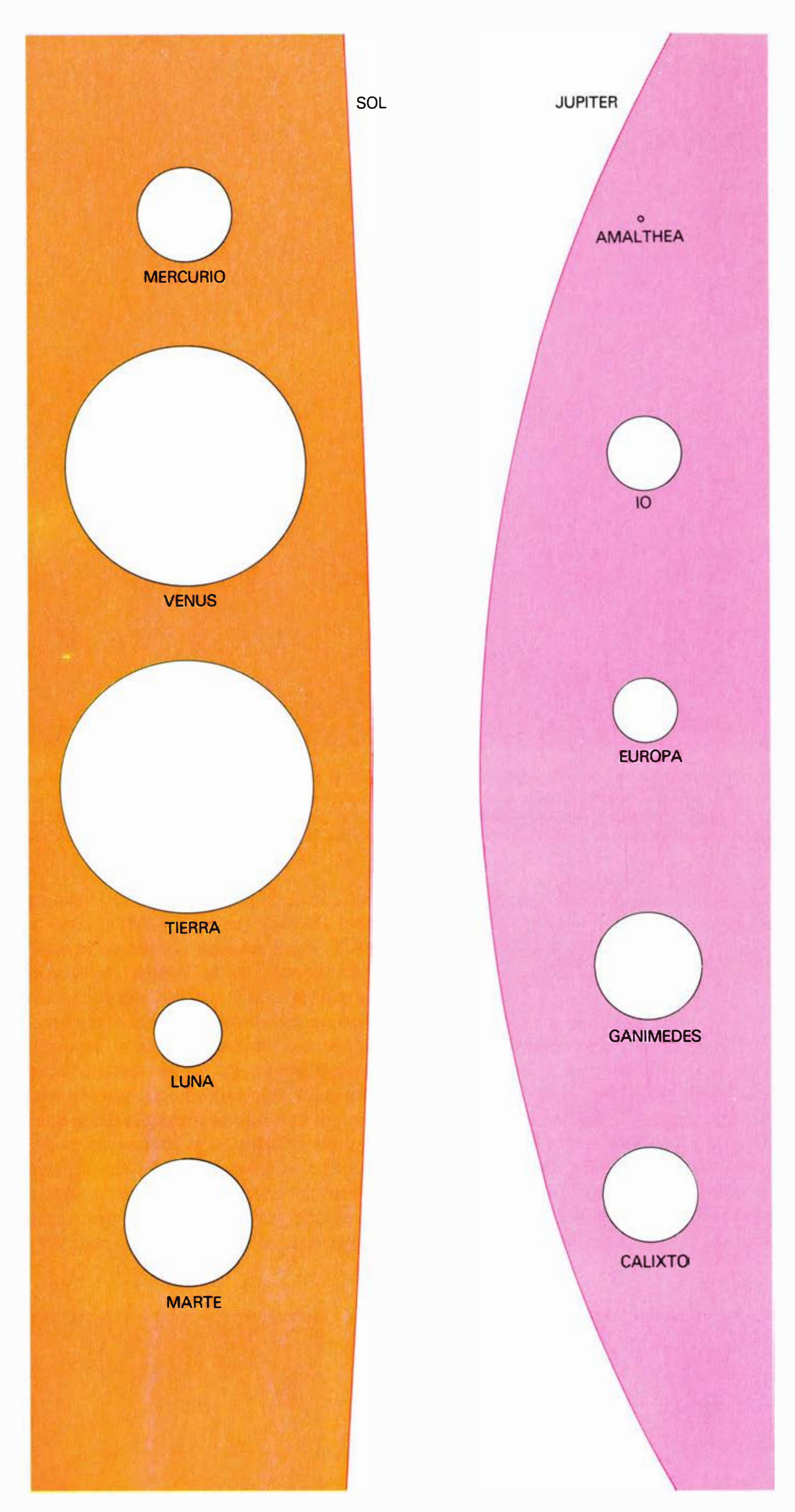
(unos 20 kilómetros por par de líneas). El corto período orbital de Io (1,77 días) también hizo posible que los dos Voyager inspeccionasen el satélite repetidamente, a muchas longitudes, durante los pocos días próximos a cada encuentro. Esto suministró un inventario y caracterización bastante completos de los penachos que señalaban las erupciones volcánicas en el satélite.

El aspecto más llamativo de los satélites de Júpiter es lo que se puede aprender de la comparación entre ellos. En su morfología está puntillosamente registrada una gran variedad de procesos geológicos y velocidades de evolución. Describiré ahora Amalthea y los cuatro satélites galileanos en orden de sus distancias a Júpiter y en relación con los descubrimientos de las misiones encomendadas a los Voyager.

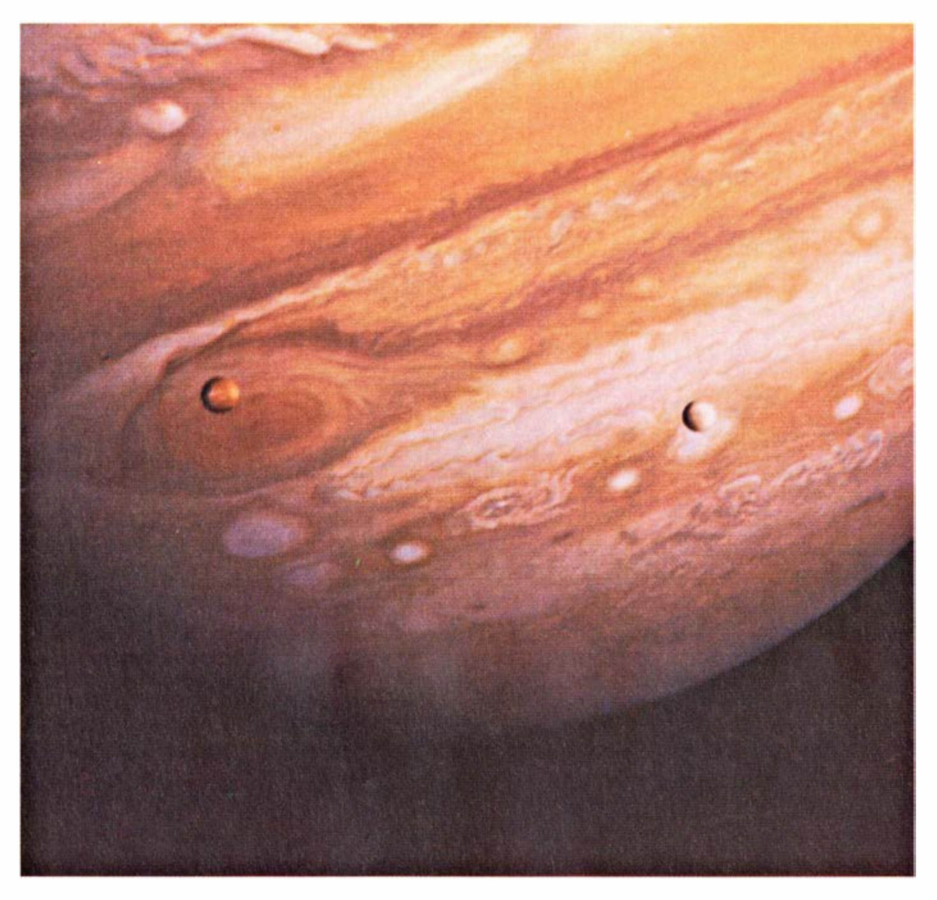
Amalthea

Amalthea pertenece a una clase especial. Aunque su tamaño es de una décima parte del de los satélites galileanos, supera en unas 10 veces los minúsculos satélites de Marte, Phobos y Deimos, que han sido exhaustivamente examinados por las misiones a Marte realizadas por vehículos espaciales. Amalthea es, por tanto, el primero de una clase de objetos de tamaño intermedio, parangonable con el de muchos asteroides, que se explora a corta distancia. De forma elipsoidal irregular, su eje mayor (de unos 270 kilómetros de longitud) apunta a Júpiter, y su eje menor (de 155 kilómetros) es perpendicular al plano de su órbita. Los datos de las imágenes del Voyager muestran que refleja la luz roja aproximadamente un 50 por ciento más que la luz violeta, lo que confirma observaciones hechas desde la Tierra. Aunque su albedo promedio, o reflectividad, oscila en torno al 5 por ciento, hay unas pocas áreas que son al menos tres veces más reflectoras. [*Véase la ilustración inferior de la página siguiente.*]

Rudolf A. Hanel, del Centro Goddard de Vuelo Espacial de la Administración Nacional de Aeronáutica y del Espacio (NASA) y sus colegas que trabajaban en el experimento infrarrojo del proyecto Voyager, descubrieron que Amalthea estaba más caliente de lo que estaría si se limitase simplemente a absorber y reemitir radiación solar y radiación reflejada de Júpiter. El exceso de calor podría ser originado por corrientes eléctricas que se desplazasen a lo largo de las líneas de fuerza del campo magnético de Júpiter o bien por el bombardeo de partículas aprisionadas en el cinturón de radiación del planeta. La forma irregular del satélite implica una rigidez interna



CUERPOS TERRESTRES **que se han estudiado de cerca en misiones espaciales son Mercurio, Venus, la Tierra, la Luna, Marte y los cuatro satélites galileanos de Júpiter. Estos cuerpos se llaman terrestres por su semejanza, en tamaño y composición, a la Tierra. Los cinco cuerpos terrestres del Sistema Solar interno se han dibujado a la misma escala a la izquierda en el orden de sus distancias al Sol. El arco de gran radio alrededor de ellos representa el borde del Sol. Los satélites galileanos se han dibujado a la misma escala junto al borde de Júpiter y en el orden de sus distancias al planeta (*a la derecha*). A esta escala, Amalthea es poco más que un punto. Desde la Tierra, poco se puede decir del satélite galileano Amalthea, excepto que es oscuro y de color rojo y que su diámetro mide escasos cientos de kilómetros.**



JUPITER, IO y EUROPA, fotografiados por el *Voyager 1* el 13 de febrero de 1979, casi un mes antes de que el vehículo espacial alcanzase su punto de máxima proximidad al planeta. La distancia era de 21,6 millones de kilómetros; la resolución es de 390 km por par de líneas de barrido de televisión. Io aparece frente al disco de Júpiter, directamente sobre la Gran Mancha Roja. Europa está a la derecha.

apreciable y un contenido bajo de sustancias volátiles.

Io

Se suponía generalmente que Io, el satélite galileano más interno, tendría una superficie antigua y salpicada de cráteres de forma muy parecida a la de la Luna terrestre. Había, sin embargo, algunos científicos que discrepaban de este punto de vista. En un sorprendente ejemplo de predicción científica, publicado justamente tres días antes del paso cercano del *Voyager 1* a Io, Stanton J. Peale, de la Universidad de California en Santa Bárbara, y Patrick M. Cassen y Ray T. Reynolds, del Centro de Investigación Ames de la NASA, indicaron que puesto que el satélite está sometido a fuerzas gravitatorias resonantes ejercidas por sus satélites hermanos, especialmente Europa, su órbita debería estar deformada respecto a la circular. Io tendría que moverse ligeramente hacia dentro y hacia fuera en el interior del poderoso campo gravitatorio de Júpiter, con el resultado de que sería repetidamente deformado por las fuerzas de marea que arrojarían a su interior una enorme cantidad de calor producida por rozamiento. Y como el calor tendría en último término que ser disipado a través de la superficie del planeta, Peale, Cassen y Reynolds especularon que "debería producirse un vulcanismo superficial extenso y recurrente".

La predicción se comprobó, pero no inmediatamente. Las primeras imágenes de Io con rasgos superficiales detectables mostraron pequeñas manchas oscuras, algunas de las cuales estaban rodeadas por débiles anillos. La primera impresión fue de que se trataba de los cráteres de impacto que se esperaban. Pero esa apreciación tuvo pronto que reconsiderarse cuando las fotografías de alta resolución no mostraron ni un solo cráter de impacto. A la máxima resolución (600 metros por par de líneas), un cráter cuyo diámetro midiera sólo un kilómetro debería ser visible. Admitiendo que en Io la velocidad de formación de cráteres de impacto mayores de un kilómetro no difiere apreciablemente de las velocidades típicas en el sistema solar interno, debía haber algún proceso en la superficie de Io que estuviera borrando los cráteres de impacto en un período de algo menos de un millón de años.

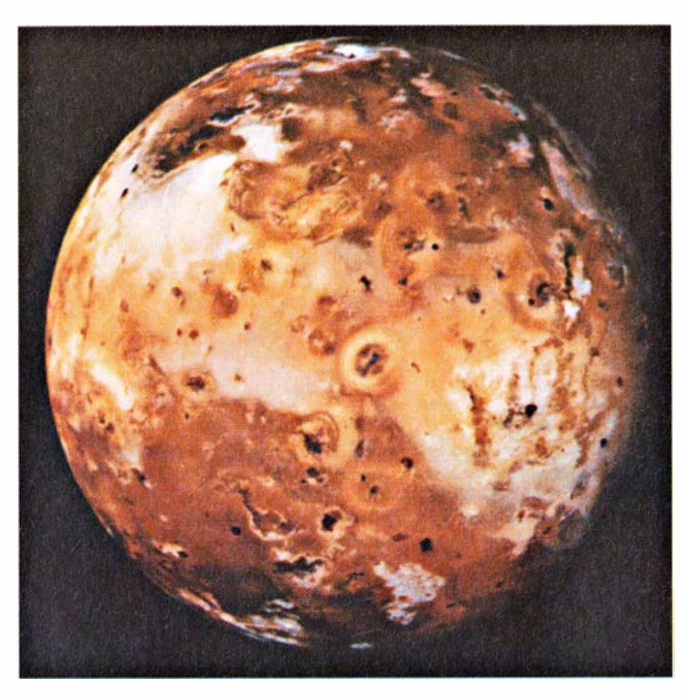
La superficie de Io muestra toda una diversidad de raras formas de suelo, con sinuosas escarpaduras y fallas [*véase la ilustración inferior de la página 65*]. Las fotografías enviadas inmediatamente antes del punto de máxima proximidad entre Io y el *Voyager 1* ofrecieron un indicio sobre el proceso que podría estar renovando la superficie del planeta. Una de tales fotografías muestra una depresión circular de un diámetro aproximado de 50 kilómetros, bordeada por un acantilado y rodeada por una formación radiante de lava [*véase la ilustra-*
ción radiante de lava [*véase la ilustra-*
ción superior de la página 64]. Esta conciones volcánicas que se observan en otros cuerpos terrestres: una caldera, o gran cráter de hundimiento, formado cuando la lava fluye hacia el exterior o es extraída de una cámara de magma, y así la superficie encima de ella se hunde. En el hemisferio que se ha visto a alta resolución se han identificado más de 100 de tales accidentes mayores de 25 kilómetros. Las corrientes que irradian desde lo que se presume constituir los centros volcánicos son de muchos colores: negro, amarillo, rojo, anaranjado y castaño. Para explicar esas observaciones, algunos miembros de los equipos Voyager han sugerido que el flujo está formado por toda una serie de sustancias que comprende desde el basalto coloreado por azufre hasta el azufre puro; incluso algunos han sugerido la existencia de lagos y océanos de azufre.

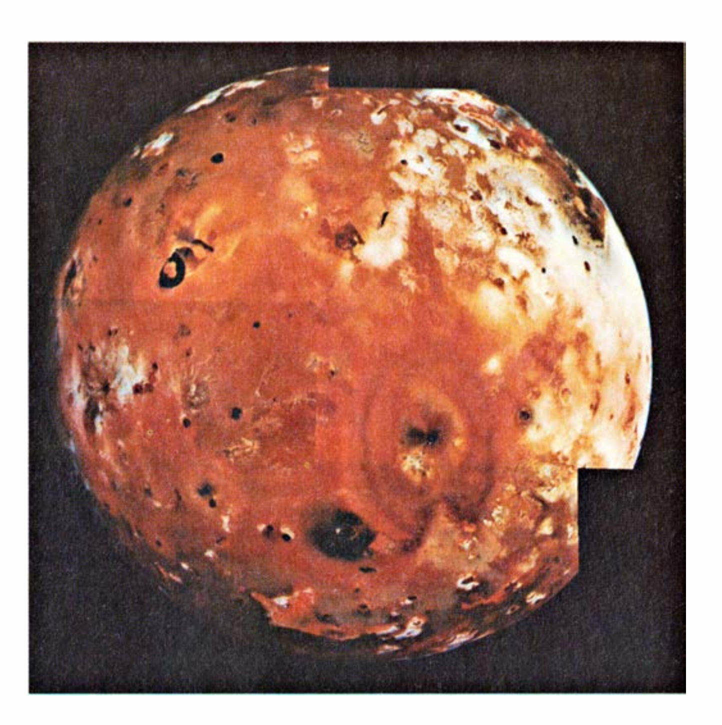


LAS MEJORES VISTAS DE AMALTHEA las registró el *Voyager 1*. Este cuerpo, semejante a un asteroide, tiene un eje mayor de unos 270 kilómetros y un eje menor de unos 155 kilómetros. La figura de la izquierda, hecha a una distancia de 695.000 km, posee una resolución de 13 km por par de líneas. La figura de la derecha, tomada a 425.000 kilómetros de distancia, tiene una resolución de ocho km.

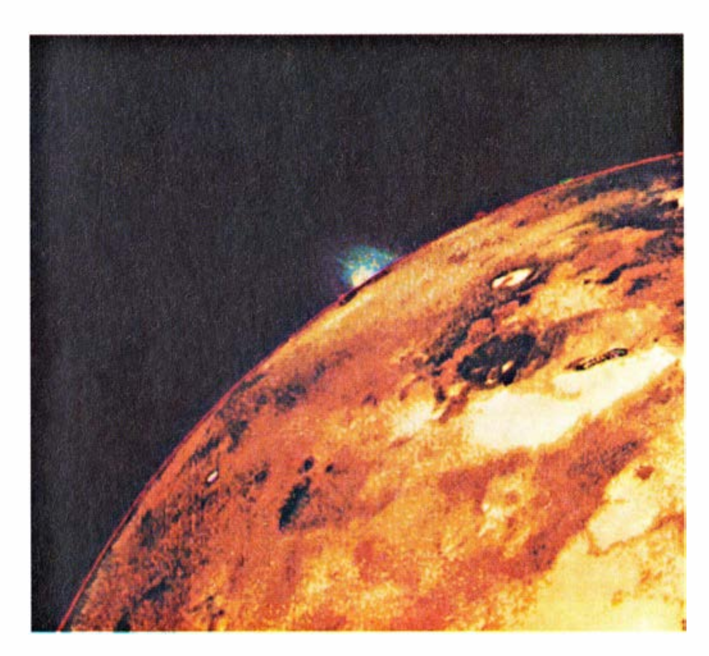
Si la superficie de Io es tan joven y la actividad volcánica tan reciente como las fotografías indican, dicha actividad debería proseguir en nuestros días. Sin embargo, quienes trabajábamos en el grupo de imágenes del Voyager nunca soñamos con detectar vulcanismo activo en Io. La probabilidad de ver un volcán activo en la Tierra con cámaras como las de los Voyager es extremadamente pequeña.

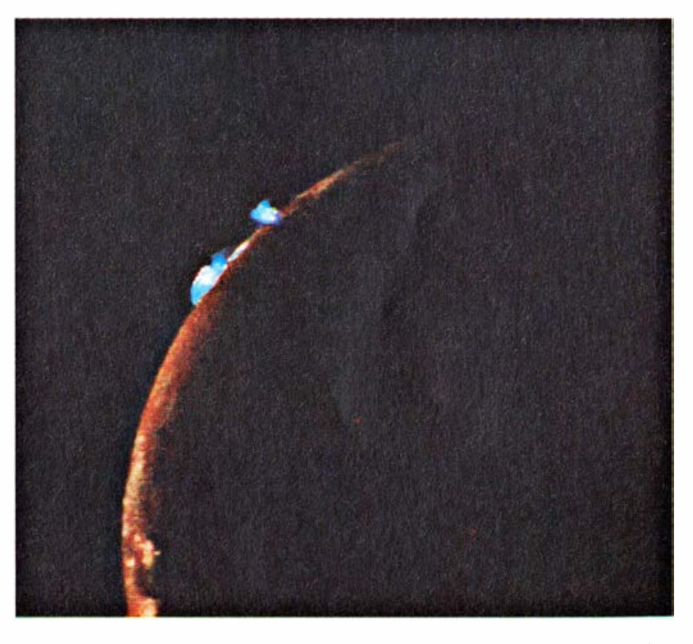
Algunos días después del encuentro del *Voyager 1* con Júpiter, Linda A. Morabito, ingeniero del Laboratorio de Propulsión a Chorro del Instituto de Tecnología de California, estaba examinando fotografías de Io tomadas a larga distancia a efectos de navegación. Se habían sobreexpuesto las películas a fin de que salieran las estrellas sitas en el campo detrás del satélite. Cuando Mora-



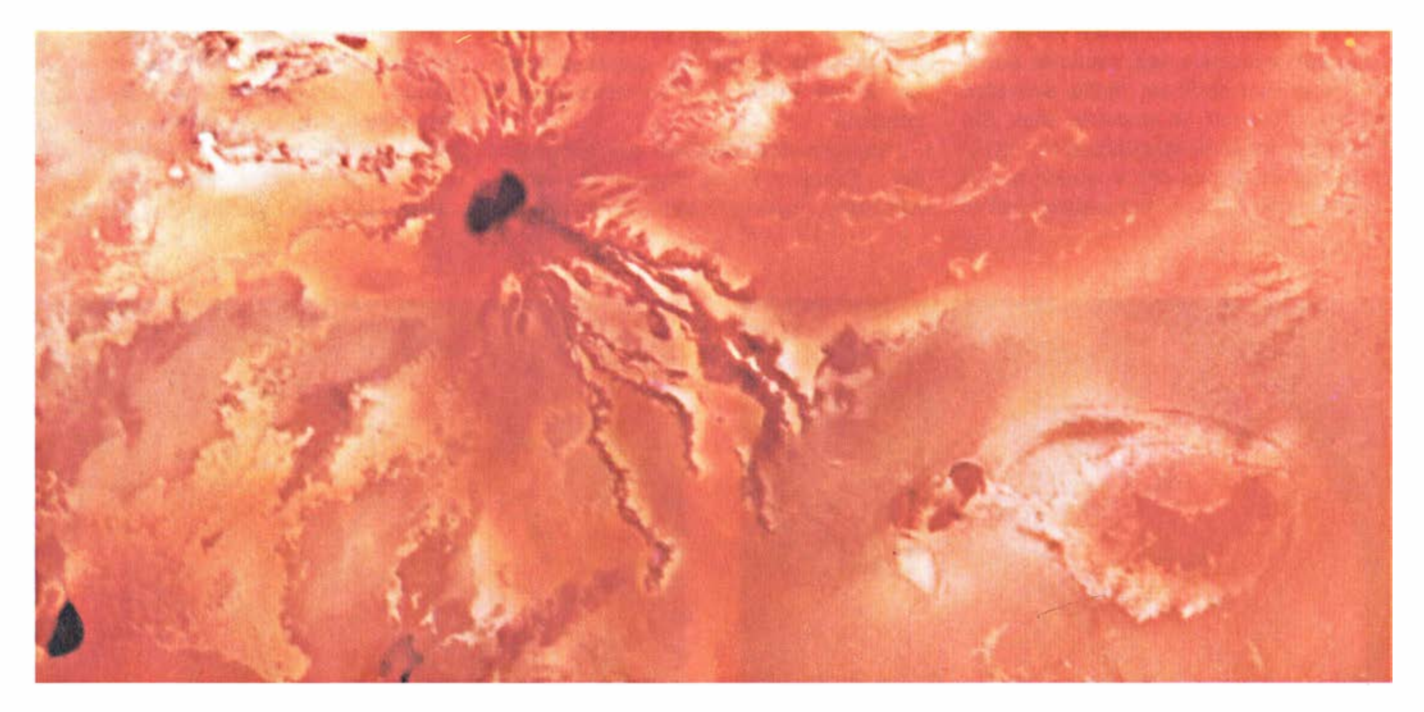


AMBAS CARAS DE IO, fotografiadas con un intervalo de tiempo de 11 horas, cuando el *Voyager 1* se acercó y después sobrepasó el satélite galileano más interno. La fotografía de la izquierda, tomada a una distancia de unos 860.000 kilómetros, muestra el hemisferio de Io que siempre está opuesto a Júpiter cuando el satélite efectúa su revolución en torno al planeta en 1,77 días. El accidente en forma de rosquilla cerca del centro es el lugar donde está situado un volcán en erupción (Penacho 3). La resolución es de unos 16 kilómetros por par de líneas. El mosaico de la derecha, de más alta resolución, se hizo en el día de máximo acercamiento del vehículo espacial y presenta al satélite galileano Io girado unos 120 grados hacia la derecha o hacia el Este. La gran región en forma de corazón en la parte más baja de la izquierda es la posición del Penacho 1, el primer volcán en erupción activa que se descubrió en el satélite. Cuatro meses más tarde, cuando llegó el *Voyager 2*, la erupción había cesado. En la ilustración de la parte inferior izquierda de la próxima página pueden contemplarse primeros planos de la misma región, tanto en fase de actividad como inactiva.





TRES PENACHOS volcánicos más en Io, del total de ocho observados, alcanzan una altura de unos 100 kilómetros en estas dos vistas. La fotografía de la izquierda, hecha por el *Voyager 1*, muestra el Penacho 2. La de la derecha, tomada por el *Voyager 2* durante una "vigilancia de volcanes" después del máximo acercamiento del vehículo espacial a Júpiter, muestra los Penachos 5 y 6. La imagen del Penacho 2 es la compañera de la imagen en la portada de este número de INVESTIGACION Y CIENCIA. La fotografía del Penacho 2 se ha construido a partir de tomas hechas en la parte visible del espectro y muestra la pequeña parte central del penacho. La fotografía de la portada incorpora una imagen hecha a través de un filtro ultravioleta (que permite el paso de longitudes de onda centradas alrededor de 3500 unidades angström) y revela que el núcleo visible está rodeado por una envoltura mayor que difunde la radiación ultravioleta. El Penacho 5 es el penacho simétrico inferior. El Penacho 6, el superior, es más pequeño y asimétrico.



RIOS DE LAVA radiantes en Io alrededor de una caldera oscura, el cráter de un volcán hundido. Esta fotografía fue tomada por el *Voyager 1* poco antes de su máximo acercamiento. La resolución es de unos dos kilómetros. En ese punto todavía no se había reconocido el primero de los penachos volcánicos. Las calderas suministraron a los investigadores la primera señal de la ausencia de cráteres de impacto que habían esperado ver mucho antes del máximo acercamiento del *Voyager 1*. Ni siquiera en las fotografías de más alta resolución, 600 metros, se pudieron observar nunca cráteres de impacto. Su ausencia indica que el vulcanismo del satélite está borrando accidentes de hasta un km de diámetro en menos de un millón de años.

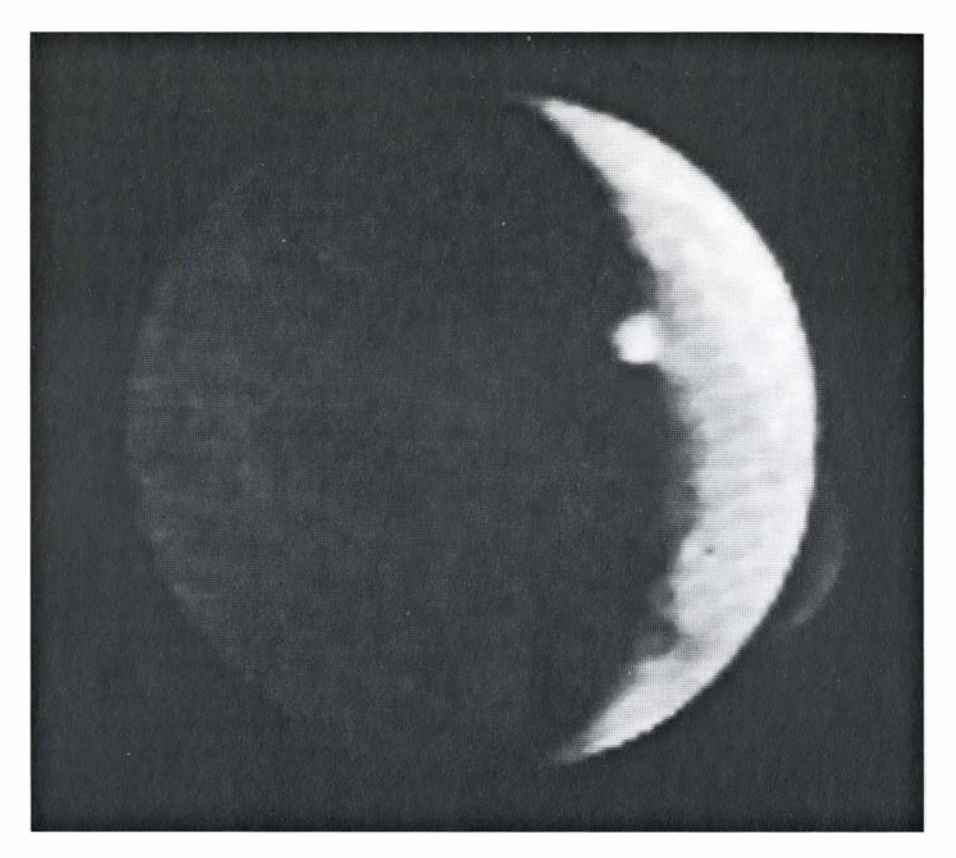
MAYOR VOLCAN ACTIVO DE IO, fuente del Penacho 1, mostrado en la imagen superior tomada por el *Voyager 1*. La figura en forma de corazón se extiende a lo largo de 1200 kilómetros en su máxima dimensión. Cuando el *Voyager 2* fotografió la misma región (*abajo*), la figura en forma de corazón se había convertido en otra configuración más simétrica. Otras imágenes del *Voyager 2* mostraron que la erupción había cesado. La imagen del *Voyager 2*, hecha a una distancia mayor, tiene una resolución de 24 kilómetros, comparada con los siete que poseían las del vehículo espacial *Voyager 1*.

TRES CALDERAS EN IO, que aparecen en fotografías tomadas con seis horas de diferencia al acercarse el *Voyager 1* al satélite desde una distancia de 374.000 kilómetros hasta 130.000 kilómetros. En la primera vista (*arriba*) los fondos de las tres calderas aparecen negros. En la segunda (*abajo*) en la caldera que está más a la derecha han aparecido manchas luminosas blanco-azuladas. Una posible explicación es que el dióxido de azufre líquido se haya filtrado desde el interior del satélite y, al alcanzar la superficie, haya explotado formando una nube de cristales de hielo y gas.

bito examinaba una de las fotografías, observó una forma grande y brillante con aspecto de sombrilla que salía del borde del hemisferio Sur de Io. Al fracasar los intentos de descartar dicha forma como un defecto fotográfico, se llegó a la conclusión de que era real. Evidentemente, una enorme nube formaba un arco hasta una distancia de 270 kilómetros de la superficie de Io. [*Véase la ilustración superior de esta página*.]

Aceptada la realidad de la nube, una búsqueda reveló que había no menos de ocho volcanes activos en Io lanzando penachos de 70 a 300 kilómetros de altura con velocidades que alcanzaban hasta un kilómetro por segundo. Algunos penachos muestran una simetría casi perfecta. En la mayoría de los casos, la materia eyectada asciende hasta unos 100 kilómetros y forma la nube a modo de sombrilla a medida que se dispersa hacia los lados y desciende de nuevo hacia la superficie. Hay penachos que muestran en el ultravioleta un globo grande y difuso y un núcleo interno en las longitudes de onda del espectro visible. [*Véase la portada y la ilustración inferior izquierda de la página anterior.*] El núcleo interno, frecuentemente asimétrico, puede estar constituido por partículas sólidas eyectadas que siguen trayectorias balísticas. La esfera ultravioleta es posiblemente la condensación de un gas que se expande simétricamente. De los ocho penachos volcánicos observados por el *Voyager 1*, siete estaban todavía allí cuando llegó el *Voyager 2* cuatro meses más tarde. Resulta, pues, evidente que los penachos se emiten de forma continua y permanecen desde varios meses hasta varios años.

Otro interesante fenómeno en Io parece ser la presencia de manchas brillantes blancas o blanco-azuladas sobre escarpaduras y fallas observadas en muchas zonas, pero particularmente en la región polar meridional. Tales características son difusas, algunas veces variables; al parecer, enmascaran la superficie subyacente, de modo que al principio se sugirió que podrían ser nubes (o materia depositada procedente de nubes) producidas por un gas que se escapaba del interior del satélite y se condensaba formando alguna clase de nieve. Una de las observaciones clave que han conducido a una explicación provisional tanto de los penachos como de las manchas blancas en las escarpaduras ha sido el descubrimiento por John Pearl, del Centro de Vuelos Espaciales Goddard, y miembros del grupo del experimento infrarrojo Voyager de que Io tiene una tenue atmósfera de dióxido de azufre. A partir de medidas de absorción en el infrarrojo se ha estimado que durante el



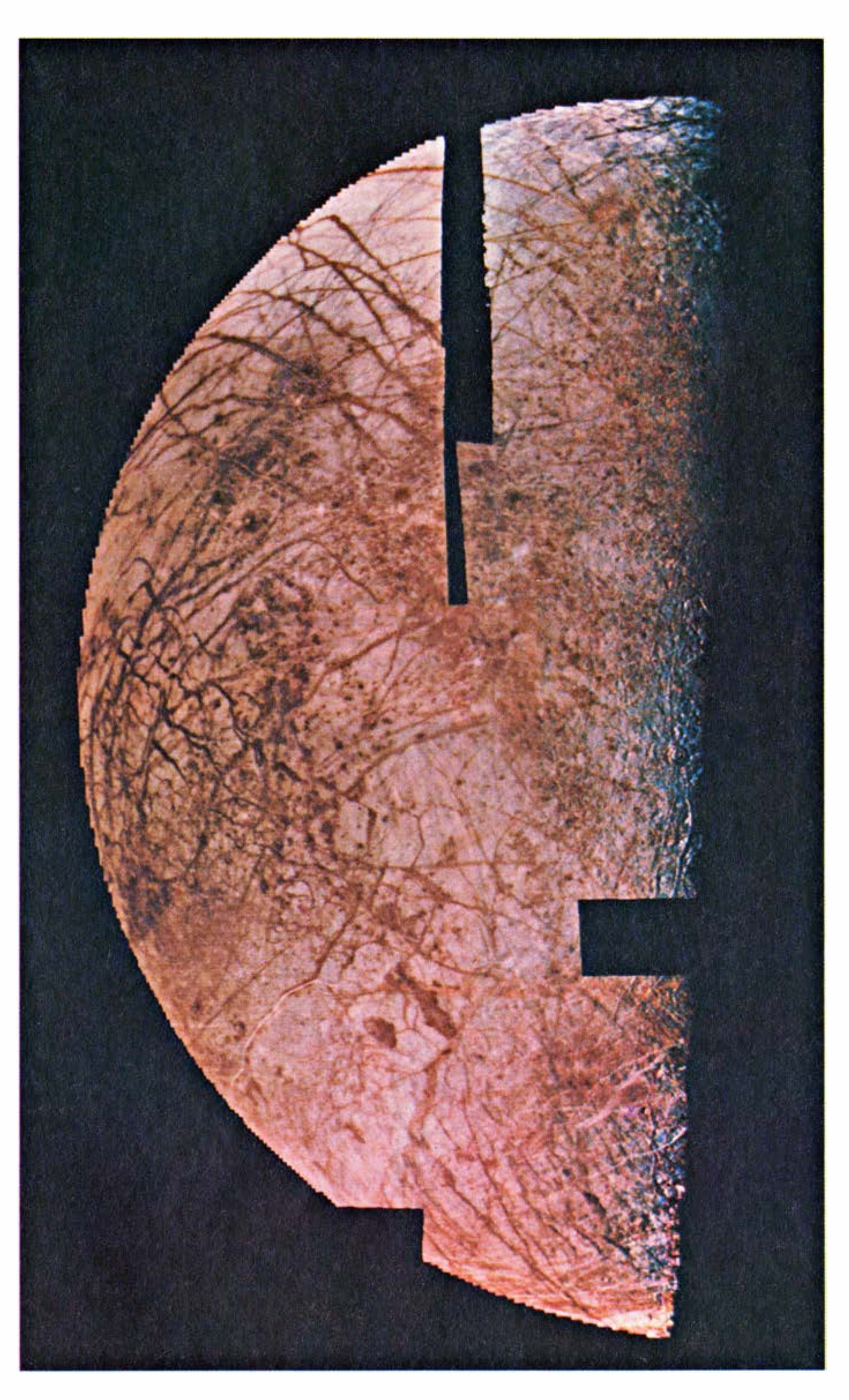
EL PRIMER VOLCAN ACTIVO EN IO fue descubierto en esta fotografía hecha por el vehículo espacial *Voyager 1* el día 8 de marzo apuntando hacia atrás desde 4,5 millones de kilómetros al satélite que había sobrepasado tres días antes. Linda A. Morabito, ingeniero del Laboratorio de Propulsión a Chorro del Cal Tech, detectó la débil nube en forma de sombrilla (Penacho 1) en el borde del satélite galileano Io cuando estaba inspeccionando una imagen procesada especialmente. Se observó entonces que la fotografía incluía un segundo penacho volcánico que había interceptado los rayos del sol naciente, creando un resplandor justamente hacia el interior de la línea del borde del hemisferio iluminado.



NUBES O DEPOSITOS SUPERFICIALES SOBRE IO visibles en la vecindad de escarpaduras en esta fotografía de una región próxima al polo Sur de dicho satélite galileano, tomada por el navío espacial *Voyager 1*. La resolución es de unos 16 kilómetros. Las manchas blancas pueden ser cristales de hielo, bien en altura o sobre la superficie, producidos cuando el dióxido de azufre líquido inmediatamente debajo de la superficie rompe a través de las fracturas en la corteza y se congela explosivamente.

día, esta atmósfera tiene una presión de alrededor de una décima de microbar (una diezmillonésima de la presión atmosférica al nivel del mar en la superficie de la Tierra).

Otra característica de la absorción en el infrarrojo, descubierta en el curso de observaciones telescópicas hechas desde la Tierra por Dale P. Cruikshank, de la Universidad de Hawai en Manoa, fue interpretada, hacia la misma época, por cierto número de observadores como posiblemente debida a escarcha de dióxido de azufre sobre la superficie del planeta. Un reciente análisis espectrofotométrico de las imágenes de los Voyager señala que las grandes regiones blancas pueden ser aproximadamente una mitad escarcha de dióxido de azufre y la otra azufre. Estos descubrimientos sugieren que el dióxido de azufre quizás abunde en las capas superficiales de Io. Si tal es el caso, el dióxido de azufre líquido sería estable en una zona cuya parte superior estaría a unos cientos de metros bajo la superficie, y se acumularía en tal zona.



EUROPA, el satélite siguiente desde Io hacia fuera, fue fotografiado por el *Voyager 2* con una resolución máxima unas ocho veces mayor que la alcanzada por el *Voyager 1*. Este mosaico de imágenes del *Voyager 2* tiene una resolución máxima de unos 4,3 kilómetros. Muestra Europa cruzada en todas direcciones por rayas y bandas que pueden representar fracturas rellenas en la helada corteza del satélite. El agua, en forma sólida y líquida, puede constituir aproximadamente el 20 por ciento de la masa de Europa. El albedo del satélite, o reflectividad de su superficie, vale casi 70 por ciento. Europa describe una revolución en torno a Júpiter cada 3,55 días a una distancia de 600.000 kilómetros.

La presencia de depósitos de dióxido de azufre líquido, de manera semejante a los acuíferos de agua líquida en la Tierra, puede proporcionar una explicación sencilla de lo que parecen ser nubes de hielo escapándose de fracturas en la corteza de Io. El dióxido de azufre fluido tendría su acceso más fácil a la superficie a lo largo de una falla o en la base de una escarpadura. Al alcanzar la superficie, quizá forzado hacia fuera por presiones artesianas, la presión que lo confinaba desciende bruscamente por debajo de un valor crítico y el líquido explota en una niebla helada que se extiende y cae de nuevo a la superficie. John F. McCauley, del Servicio de Reconocimiento Geológico de los Estados Unidos, ha supuesto que este proceso actúa realmente erosionando materia a lo largo de las pendientes.

También se ha hecho uso del "acuífero" de dióxido de azufre para explicar los penachos volcánicos, mucho mayores. Bradford A. Smith, de la Universidad de Arizona, y sus colegas del grupo encargado de estudiar las imágenes del proyecto Voyager han indicado que el dióxido de azufre líquido puede causar violentas erupciones volcánicas de una manera muy parecida a como lo hace el agua en la superficie de la Tierra. En el modelo propuesto para Io, el calor, generado por mareas en la porción de litosfera constituida por silicatos, es transportado hacia arriba por el azufre fundido que ha estado en contacto con rocas muy calientes. Cuando el azufre fundido entra en contacto con el dióxido de azufre líquido, la mezcla comienza a ascender hacia la chimenea volcánica y se expande rápidamente al descender la presión, y el dióxido de azufre se vaporiza. Cuando alcanza la superficie, su velocidad estará entre 500 y 1000 metros por segundo, lo que concuerda con las velocidades obtenidas para penachos que ascienden de 100 a 300 kilómetros por encima de la superficie. Este modelo requiere grandes depósitos separados de azufre fundido, tal vez lo bastante grandes para llamarles océanos, bajo la superficie de Io. Algunas de las estrías volcánicas observadas pueden por tanto estar constituidas por azufre casi puro, cosa que también ha sido propuesta por Carl Sagan, de la Universidad Cornell.

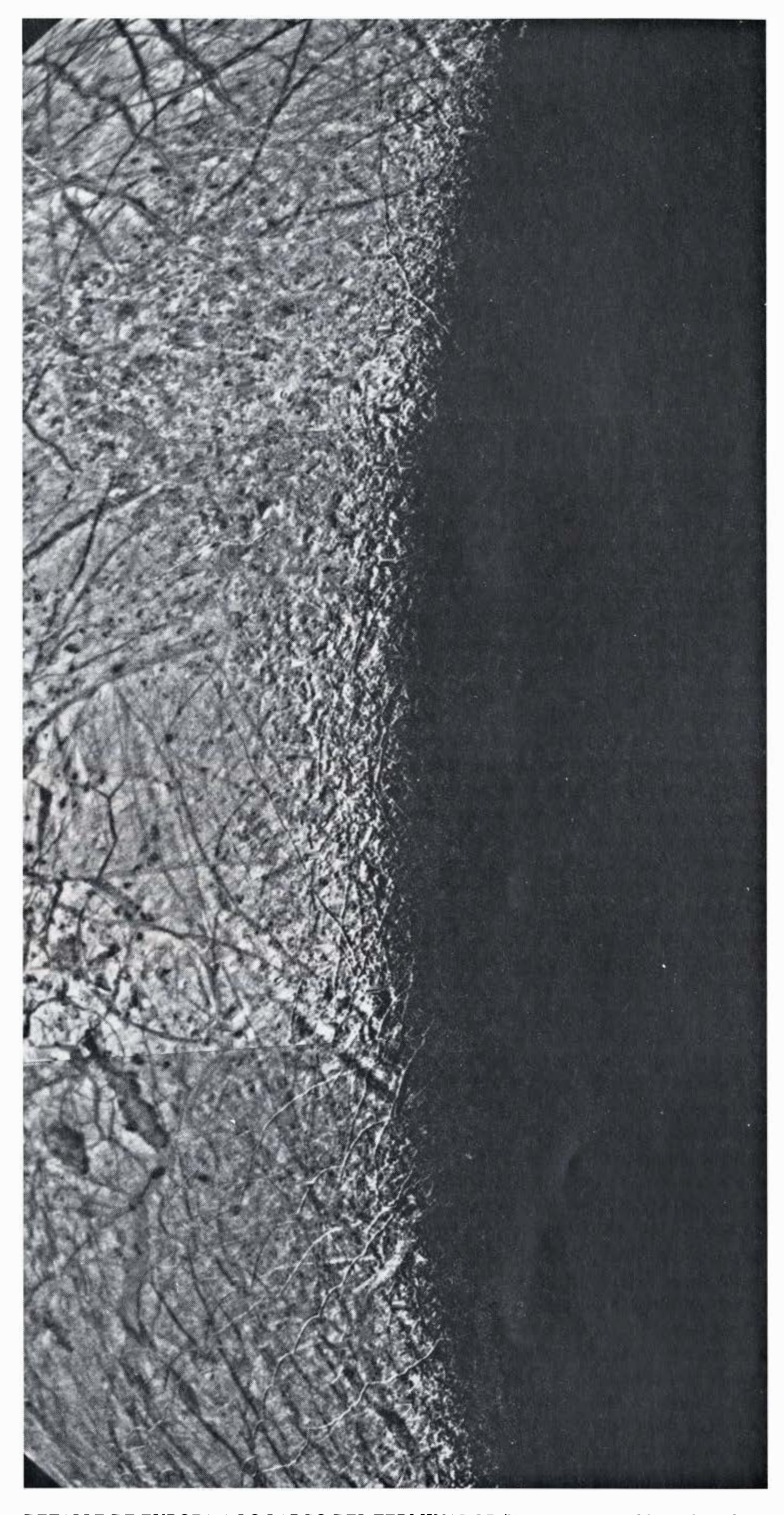
En el otro extremo están los modelos propuestos por Michael H. Carr y Harold Masursky, de la Inspección Geológica de los Estados Unidos, en los cuales el azufre y los compuestos de azufre son simplemente agentes colorantes de los ríos de lava del vulcanismo silíceo corriente, tal como el de la Tierra, y quizás actúen distintos mecanismos en diferentes partes del satélite. Hay acuerdo general en que los vívidos colores que se ven en la superficie están en concordancia con la amplia gama de formas moleculares del azufre que se sabe son estables a las temperaturas de la superficie de Io, que varían entre 60 y 120 grados K.

Europa

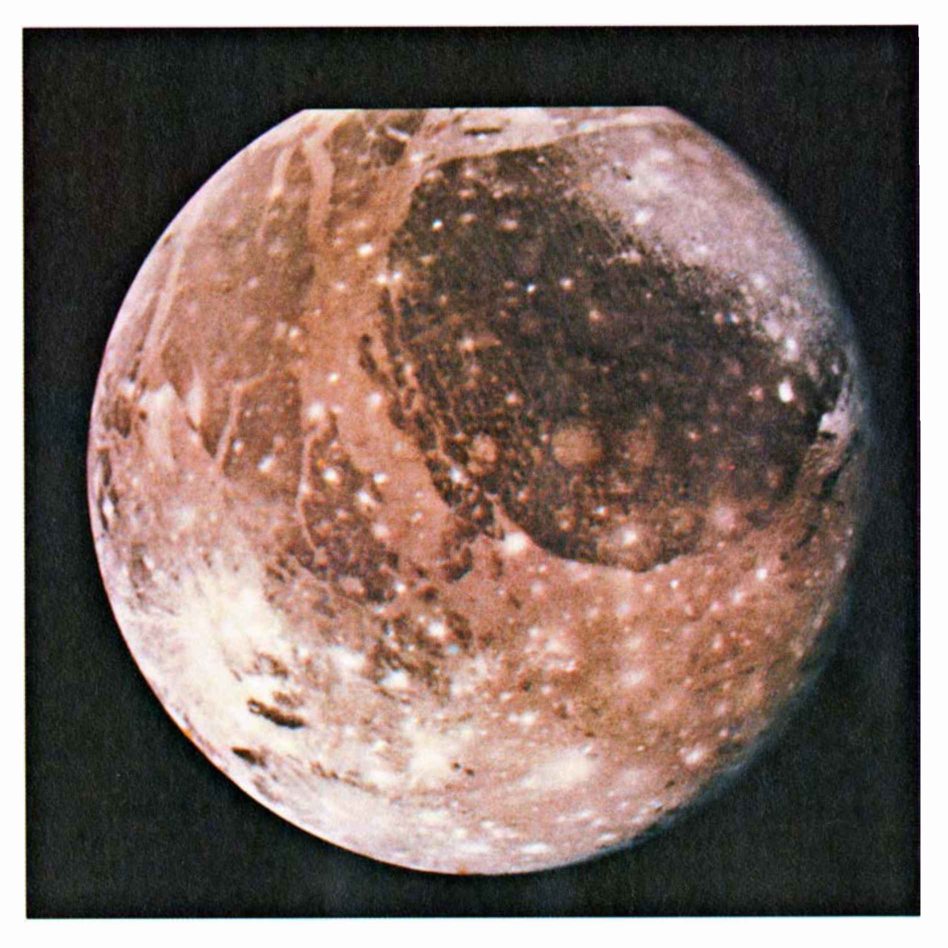
Europa era el satélite galileano más alejado de la trayectoria del *Voyager 1* (la distancia mínima entre ambos fue de 732.270 kilómetros); de ahí que se le fotografiara con la resolución más baja, de alrededor de 33 kilómetros por par de líneas en las mejores fotografías. Esta resolución es comparable a la de una fotografía de la Luna terrestre que ocupase un tercio de la altura de una pantalla de televisión corriente. El *Voyager 2*, que se acercó hasta 204.000 kilómetros, suministró una resolución ocho veces mejor. Las imágenes de Europa del *Voyager 1* muestran un cuerpo casi blanco, con rasgos de carácter global y configuraciones que son desvaídas y de débil contraste. Su región ecuatorial muestra dos tipos básicos de terreno: regiones más oscuras, moteadas, y regiones más claras; ambas clases, atravesadas por una serie de franjas oscuras de unas decenas de kilómetros de anchura que se extienden en algunos casos a lo largo de miles de kilómetros.

En las imágenes del *Voyager 2*, de mayor resolución, las franjas aparecen como una vasta confusión de líneas que se cortan [*véase la ilustración de la página precedente*]. El terreno oscuro moteado se resuelve en una serie de depresiones y mesas interconectadas cuyas dimensiones típicas son de unos pocos kilómetros. Muchas de las depresiones pueden ser cráteres de impacto, pero su identificación se hace incierta.

Sólo se han reconocido tres probables cráteres de impacto. Tienen diámetros entre 18 y 25 kilómetros, y despliegan morfologías ampliamente diferentes. Hay un cráter reciente y en forma de tazón; otro es poco profundo y rodeado de un sistema de rayos oscuros. El tercero parece estar elevado sobre un pedestal, como si se hubiera excavado la región circundante. En algunas imágenes aparecen signos de numerosos cráteres pequeños de unos pocos kilómetros



DETALLE DE EUROPA A LO LARGO DEL TERMINADOR (línea que separa el hemisferio iluminado de oscuro) visto a alta resolución (unos 3,8 kilómetros por par de líneas) en este mosaico de imágenes tomadas por el vehículo espacial *Voyager 1*. La baja inclinación de la luz que lo ilumina hace posible estimar que las largas y estrechas crestas se elevan sólo unos 100 metros sobre la superficie. Las crestas son más abundantes en el hemisferio Sur de este satélite galileano. La mayor parte del terreno en Europa está constituido por depresiones y mesas entrelazadas. Aunque muchas de las depresiones podrían ser cráteres de impacto, resultan demasiado pequeñas para identificarlas de tales.



GANIMEDES es el mayor de los satélites de Júpiter. Esta fotografía, que fue tomada por el *Voyager 2* a una distancia de 1.217.000 kilómetros, muestra el hemisferio del satélite que siempre está opuesto a Júpiter. El hemisferio está dominado por una inmensa área oscura, el mayor de los restos de una antigua corteza salpicada de impactos de meteoritos. En su historia temprana, Ganimedes pudo haber estado tan densamente cubierto de cráteres como lo está hoy Calixto. Ganimedes tiene una densidad de sólo dos tercios de la de Europa y por tanto puede tener una fracción mayor de agua, probablemente un cincuenta por ciento. Ganimedes y Calixto, los dos satélites galileanos más exteriores, poseen propiedades a gran escala que resultan ser similares y sus superficies evolucionaron de forma paralela.



HEMISFERIO DE GANIMEDES ENFRENTADO A JUPITER, fotografiado a alta resolución por el *Voyager 1*. El área aquí mostrada tiene una dimensión transversal de unos 1600 kilómetros y tiene una resolución máxima de unos cuatro kilómetros por par de líneas. La corteza de Ganimedes está formada presumiblemente en su mayor parte por hielo, que es más oscuro en las regiones más antiguas. Alrededor de los cráteres con rayos, el hielo puede ser blanco, habida cuenta de su mayor nitidez.

de diámetro, particularmente a lo largo de la zona terminal (la zona al borde del lado iluminado del satélite). La superficie de Europa tiene una edad de al menos cientos de millones de años y quizá de miles de millones de años.

La ausencia de relieve importante a lo largo de la zona terminal, que es donde sería más visible, sugiere que el agua, el más probable, y principal, componente volátil del satélite, ha ascendido desde el interior a la superficie y ha formado una espesa capa de hielo que enmascara el relieve topográfico. Una capa de 100 kilómetros de espesor sería ciertamente lo bastante espesa para velar cualquier relieve que pudiera existir en la litosfera silícea.

Igual que ocurrió con todos los otros cuerpos planetarios del sistema solar, Europa se creó presumiblemente por reunión y soldadura de materiales en órbita alrededor del Sol, hace cuatro o cinco mil millones de años. Después, de nuevo al igual que los demás cuerpos, continuó sometido a un intenso bombardeo de meteoritos, grandes y pequeños, durante otros quinientos millones de años quizá. Su apariencia actual es probablemente un reflejo exacto de la relación entre su intenso bombardeo al principio y su ulterior historia térmica. Si su corteza de hielo se hubiera formado, congelado y hecho rígida en el primitivo período de intenso bombardeo, las huellas de los impactos serían visibles con claridad hoy día. Resulta, pues, evidente que la corteza permaneció caliente, blanda y fluida hasta una época suficientemente tardía en su historia para enmascarar las pruebas del bombardeo.

Se han propuesto dos mecanismos para explicar cómo pudo mantenerse la temperatura de la superficie del satélite durante este período primitivo. Fraser P. Fanale y sus colaboradores del Laboratorio de Propulsión a Chorro propusieron, hace unos años, la hipótesis de que el calor liberado en la desintegración de elementos radiactivos habría sido por sí mismo suficiente para mantener blando, y quizá líquido, el recubrimiento de hielo de Europa a profundidades por debajo de unas decenas de kilómetros. Recientemente, Cassen y sus colegas han sugerido la siguiente alternativa: la misma clase de rozamiento debido a las mareas que calienta Io calentaría Europa, aunque en menor grado. También sugieren que las franjas oscuras de la superficie del satélite podrían ser grietas rellenas, resultantes de la simple dilatación cuando los océanos se congelaron. La hipótesis, sin embargo, quizá no pueda dar cuenta del aumento de área de 10 a 15 por ciento que implica la

anchura de las grandes configuraciones oscuras. Hay otros muchos modelos posibles.

Ganimedes y Calixto

Ganimedes y Calixto, los dos satélites galileanos más exteriores, se tratarán aquí como una pareja porque sus propiedades en gran escala son similares y porque las pruebas aportadas por los Voyager indican que sus superficies evolucionaron, hasta cierto punto, de una manera paralela. Aunque Calixto es el más distante de Júpiter, lo consideraré primero por su mayor inteligibilidad y porque las diversas fases de su desarrollo parecen evidenciarse en algunos de los terrenos más antiguos de Ganimedes.

El hemisferio de Calixto que mira a Júpiter se reprodujo a alta resolución (entre 2,3 y siete kilómetros por par de líneas) cuando el *Voyager 1* pasó junto al satélite a una distancia de 124.000 kilómetros [*véase la ilustración de la página 59*]. La superficie está casi saturada de cráteres; sin embargo, Calixto tiene un aspecto completamente diferente de la Luna terrestre.

Por ejemplo, hay una completa ausencia de relieve visible en el borde brillante de Calixto. La mitad izquierda del hemisferio está dominada por un sistema de anillos concéntricos centrados en una brillante región circular a 10 grados Norte del ecuador y de unos 600 kilómetros de diámetro. Los anillos, separados entre sí de 50 a 200 kilómetros, se extienden hasta un radio de unos 1500 kilómetros.

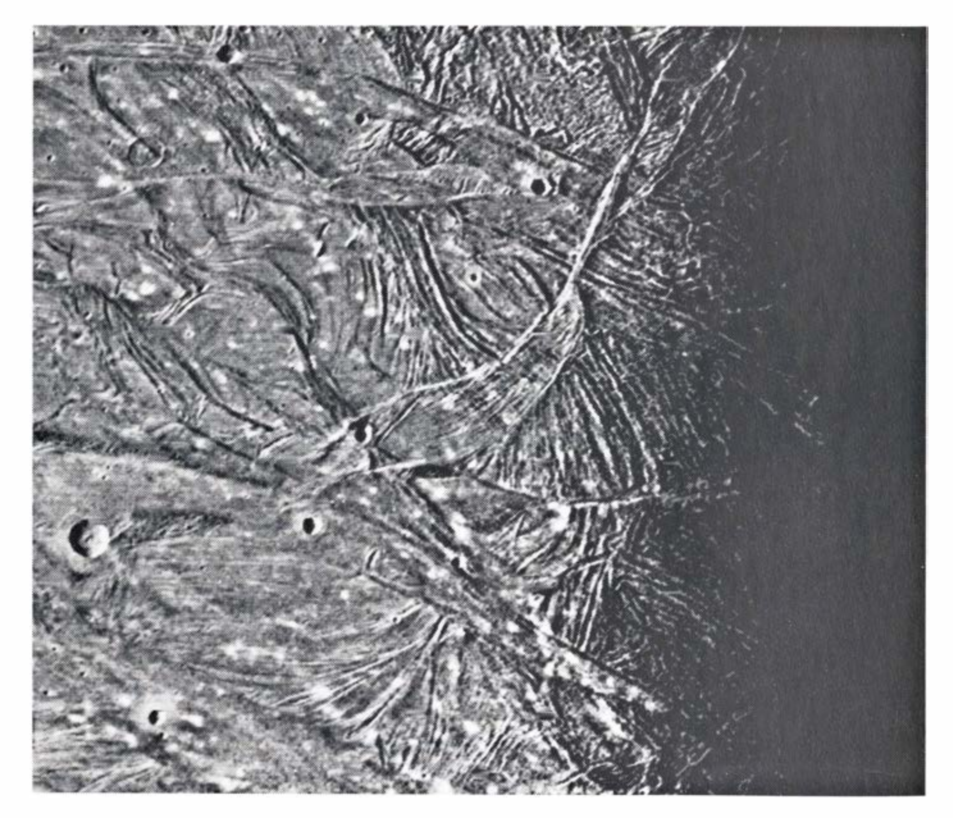
La población de cráteres dentro de la parte interior del sistema de anillos es alrededor de una tercera parte de la existente en el resto de Calixto. La población en la parte exterior del sistema posee una densidad similar a la de las otras regiones, pero a juzgar por los cortes e intersecciones de los accidentes, alrededor de dos tercios de los cráteres se formaron antes que el sistema de anillos. Evidentemente, los cráteres preexistentes dentro de unos 350 kilómetros de la zona central desaparecieron, mientras que los cráteres a mayores distancias se conservaron. Los cráteres jóvenes están superpuestos tanto en los anillos como en la zona central. La explicación más probable es que el gran impacto que creó los anillos tuvo lugar a principios de la historia del satélite, cuando la corteza no era suficientemente rígida para soportar y retener las formas topográficas normalmente asociadas a las grandes depresiones de impacto.

Los anillos mismos pudieron haberse formado dinámicamente como consecuencia inmediata del impacto de un gran cuerpo, o pudieron haber constituido una respuesta retardada al impacto, al tratar la región central de recuperar su posición y reajustarse la superficie circundante. Más tarde, la corteza de hielo se hizo rígida al congelarse hasta profundidades considerables. Como veremos, una configuración similar de tamaño comparable se ha conservado parcialmente en la zona más antigua de la corteza de Ganimedes.

Las fotografías de Ganimedes hechas por el *Voyager 1* durante su aproximación al satélite muestran un cuerpo que se parece notablemente a la Luna terrestre. Una configuración visible a esta escala que no se parece a nada observable de la Luna, sin embargo, es una compleja red de bandas brillantes irregulares, lineales, interceptadas y bifurcadas que cruzan en todas direcciones el disco del satélite. A dicha escala, las configuraciones guardan muy poca semejanza con los rayos de los cráteres de impacto repartidos sobre el resto de la superficie. La hipótesis de partida fue que las configuraciones podrían ser tectónicas.

Al acercarse más el *Voyager 1* a Ganimedes, se pudieron reconocer dos tipos básicos de terreno: regiones poligonales densamente salpicadas de cráteres, de dimensiones de hasta unas decenas de kilómetros, rodeadas por regiones más jóvenes de terreno estriado. Se encontró que el terreno estriado ocupaba la parte que, a baja resolución, se veía como configuraciones brillantes y lineales. El terreno estriado está constituido por surcos y lomas paralelos y estrechamente espaciados, cada uno de 5 a 15 kilómetros de altura y de hasta varios cientos de kilómetros de longitud. En algunas regiones se pueden contar hasta 20 surcos y lomas paralelos.

La densidad de los cráteres en el terreno estriado de Ganimedes es extremadamente variable, alcanzando desde una densidad equivalente a la que se encuentra en el terreno cubierto de cráteres hasta una densidad del orden de la décima parte. La conclusión es que la formación del terreno estriado comenzó en la historia temprana del satélite y continuó durante un largo tiempo que comprendía el período de bombardeo intenso. La mayor parte del terreno estriado es un mosaico de sistemas discretos, con las estrías de un sistema terminando bruscamente en la frontera del sistema adyacente. En algunos casos, sin embargo, diversos sistemas de estrías se interfieren entre sí. En tales casos, el sistema más antiguo no queda borrado por el otro, lo que implica que el proceso que forma las estrías no lleva consigo movimiento lateral. Se ha sugerido que las estrías se han formado mediante un



TERRENO SURCADO Y RETORCIDO, rasgo característico de Ganimedes. Esta vista del *Voyager 1* a lo largo del terminador, desde una distancia de 135.000 kilómetros, tiene una resolución de unos 2,4 kilómetros. Las crestas y surcos paralelos y estrechamente espaciados alcanzan una longitud transversal de entre 5 y 15 kilómetros. Se pueden ver sistemas más jóvenes de surcos cortando a los sistemas más antiguos. En la zona que aparece aquí, el terreno surcado ha reemplazado a la corteza antigua.

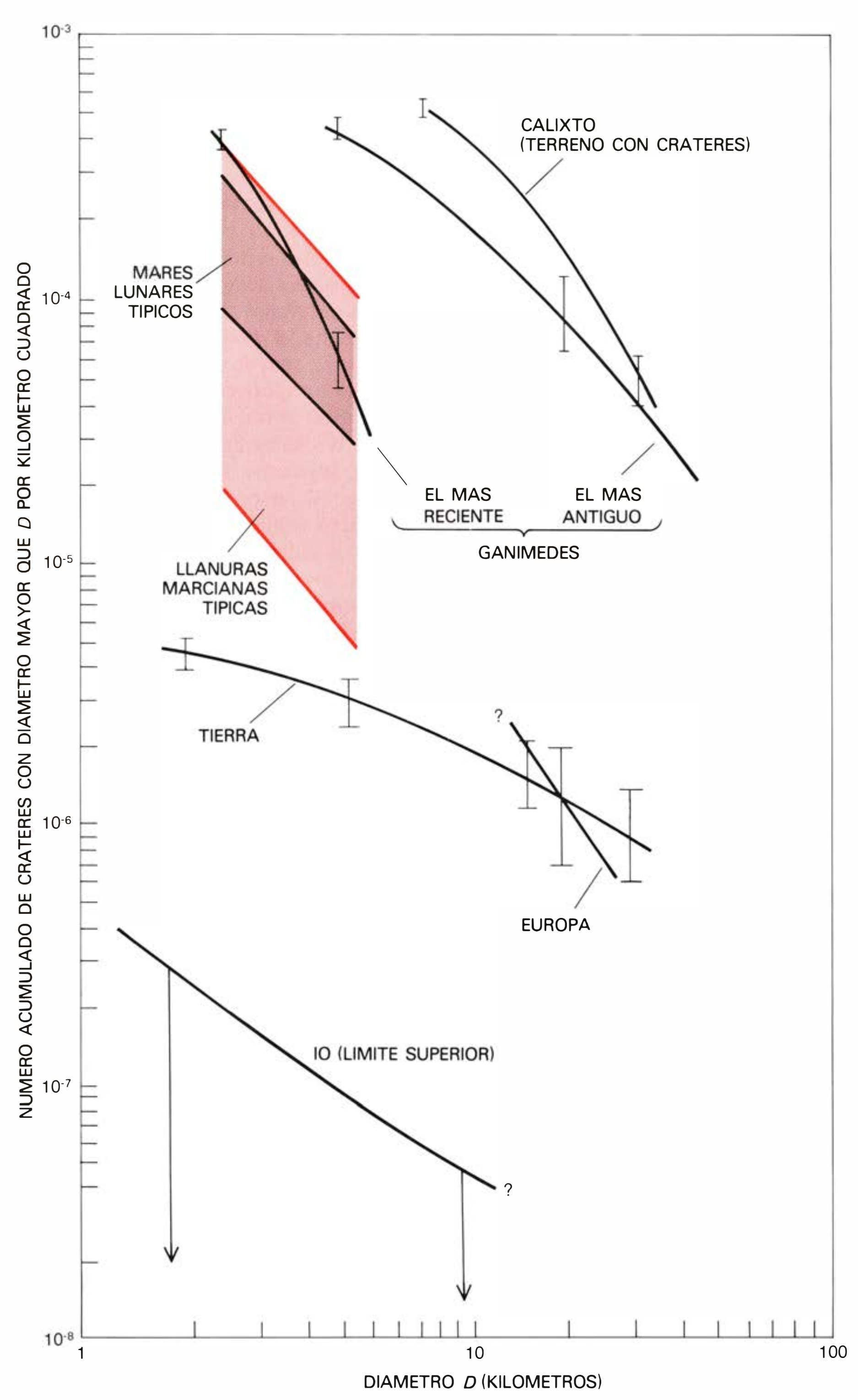
proceso semejante al que crea los diques intrusivos en la corteza terrestre: materia fluida podría inyectarse en las fallas y congelarse después.

Los pases de ambos Voyager cerca de Ganimedes se complementaron muy bien entre sí. El *Voyager 1* exploró el hemisferio del satélite que mira a Júpiter; el *Voyager 2* encontró al satélite antes de que el vehículo espacial alcanzara su punto más próximo a Júpiter y, por tanto, consiguió una cobertura en alta resolución del hemisferio opuesto [*véase la ilustración superior de la página 68*]. La característica más llamativa de la vista de Ganimedes obtenida por el *Voyager 2* es una gran región circular de terreno antiguo, oscuro y salpicado de cráteres que cubre aproximadamente un tercio de la totalidad del hemisferio. Es la única característica que se puede identificar fácilmente con telescopios en la Tierra. Atravesando la región oscura hay una serie de vetas gigantes y brillantes que son paralelas y están ligeramente curvadas. En las imágenes de alta resolución presentan una sorprendente semejanza con las estructuras en anillo de Calixto. Aunque diversos estudios han mostrado que las rayas de Ganimedes forman círculos concéntricos, no hay rastro en su centro de simetría de una zona brillante y destrozada por el impacto como la existente en Calixto. Resulta evidente que la formación de terrenos más jóvenes, principalmente los estriados, ha borrado la historia de un antiguo punto de impacto.

Aunque los sistemas de anillos de Calixto y Ganimedes presentan marcadas semejanzas, donde los anillos de Calixto tienen la forma de lomas con cima plana, los de Ganimedes son arrugados. Estas arrugas sugieren el aspecto que tenían las lomas de Calixto antes de ser rellenadas. Tales diferencias pueden reflejar variaciones en los detalles de la historia de ambos cuerpos, incluyendo diferencias en la resistencia de la primitiva corteza, diferencias en la cronología del calentamiento y enfriamiento y diferencias en la disponibilidad de materiales fluidos capaces de aflorar a la superficie.

Otra importante observación es la amplia variación en el aspecto de los cráteres de impacto de Ganimedes. Algunos cráteres manifiestan un aspecto joven y reciente. En el otro extremo hay formas circulares que son poco más que "fantasmas" o manchas en la antigua corteza. Evidentemente, los cráteres fantasmas, más brillantes, se formaron cuando la corteza de Ganimedes estaba demasiado caliente y blanda para registrar de un modo permanente el impacto que los originó. Un examen detallado indica que cuanto más joven o pequeño es el cráter, en mejor estado de conservación se encuentra su relieve. Eugene M. Shoemaker y sus colegas del Servicio de Inspección Geológica de los Estados Unidos han llegado a la conclusión de que el registro de la formación de cráteres en los terrenos antiguos de Ganimedes y Calixto es también un registro del gradiente térmico y de la resistencia de la corteza superior en función del tiempo a medida que los dos cuerpos comenzaron a enfriarse.

El descubrimiento por el *Voyager 2* de una depresión de impacto grande y reciente cerca del polo Sur de Ganime-



EL TAMAÑO Y FRECUENCIA DE LOS CRATERES DE IMPACTO en Europa, Ganimedes y Calixto cae dentro del dominio medido en la Tierra, la Luna terrestre y Marte. Sin embargo, Io carece de cráteres de impacto superiores a la máxima resolución de las imágenes Voyager sobre el satélite: alrededor de un kilómetro. La curva correspondiente a Europa es muy incierta ya que solamente se han podido identificar de manera fidedigna tres cráteres en las fotografías de los Voyager. Las estimas de cráteres de impacto en la Tierra se deben a Richard Grieve y Michael R. Dence, del Departamento de Energía del Canadá. Cuanto más antigua sea la superficie de un planeta, mayor su número de cráteres.

des es una prueba más de importantes cambios acontecidos en la resistencia y rigidez de la corteza del satélite. La depresión se parece a algunas de las más importantes depresiones de impacto en el sistema solar interno. En tomas a alta resolución podemos apreciar que la depresión, de fondo plano, está rodeada de macizos en forma de bloques y una capa de materia expulsada; ello indica que hacia la época en que se formó la depresión, la corteza se había endurecido lo suficiente para soportar relieve montañoso en gran escala. El hecho de que haya numerosos cráteres grandes superpuestos a la depresión sugiere que ésta se formó en los primeros mil millones de años después de la constitución de Ganimedes, cuando el satélite estaba todavía sometido al bombardeo de cuerpos de tamaño considerable. Queda claro que la corteza se estaba enfriando rápidamente y congelándose durante los primeros cientos de millones de años de la existencia del satélite.

A partir de la comparación de Ganimedes con Calixto, podemos reconstruir, a grandes rasgos, la historia de la evolución de Ganimedes. Se puede imaginar que la antigua superficie del satélite era blanda, helada y oscura, parecida a la actual superficie de Calixto. El gran conjunto de rayas paralelas y ligeramente curvadas en el hemisferio Norte se formó cuando un gran meteorito perforó la blanda superficie y originó oscilaciones y esfuerzos que fracturaron la corteza en una enorme configuración de anillos concéntricos. El agua llenó las fracturas y se congeló. Más tarde se formaron los casquetes polares, y el terreno estriado comenzó lentamente a desarrollarse y crecer. A medida que secciones de la corteza deslizaban unas junto a las otras, se formaron fallas en ángulo respecto a las líneas de deslizamiento que proveyeron los contrapesos en las líneas. Después, la corteza se hizo lo suficientemente rígida para conservar un registro duradero de los impactos de meteoritos. La joven depresión de impacto cerca del polo Sur y un gran cráter con brillantes rayos cerca del ecuador fueron las últimas configuraciones en gran escala creadas en el hemisferio opuesto a Júpiter que fue fotografiado por el *Voyager 2*. En los varios miles de millones de años transcurridos desde esos tempranos sucesos en la evolución de Ganimedes, el satélite ha permanecido evidentemente tranquilo.

La historia archivada en los cráteres

Uno de los principales objetivos de la ciencia planetaria es comparar las historias geológicas de los planetas terrestres y de los satélites del sistema solar, tal como la revelan las imágenes detalladas enviadas por los vehículos espaciales. Para establecer las escalas de tiempo relativas de esas historias, a partir de las imágenes, el método principal seguido consiste en determinar el registro de la formación de cráteres de impacto de acuerdo con una hipótesis sencilla: cuanto más antigua es la superficie, tanto mayor será el número de cráteres sobre ella. Mucho más difícil resulta establecer escalas de tiempo absolutas a partir del estudio de los cráteres, porque ello requiere el conocimiento de la velocidad de formación de cráteres para cada uno de los cuerpos.

Se dispone ya de un amplio conjunto de pruebas sobre la manera como la velocidad de formación de cráteres difería de una a otra parte del sistema solar. Los modelos de la velocidad a la cual los asteroides fueron creando cráteres (y en menor grado, los cometas), junto con observaciones del número de cráteres en distintos cuerpos, indican que las velocidades medias a lo largo de los últimos miles de millones de años eran similares para Mercurio, Marte, la Tierra y la Luna dentro de un factor de quizá dos.

Sin embargo, Shoemaker ha indicado que a lo largo de los últimos, pocos, miles de millones de años, las velocidades de formación de cráteres para los satélites de Júpiter pudieron ser notablemente menores que la del Sistema Solar interno, quizás entre 10 y 100 veces menores. Su razonamiento es el siguiente. Aunque los asteroides salgan desviados fuera del Sistema Solar interno hacia la región de la órbita de Júpiter, probablemente pronto volverán a ser desviados por el poderoso campo gravitatorio de Júpiter hacia otras partes del Sistema Solar. Por tanto, sólo tienen una pequeña probabilidad de chocar con Júpiter o sus satélites. Siendo éste el caso, el flujo de objetos menores en el sistema joviano está probablemente dominado por los cometas periódicos, que son quizá diez veces menos abundantes que los asteroides y tal vez hasta 100 veces.

Otro factor que influye en la velocidad de formación de cráteres en los satélites es la enorme masa de Júpiter, que aumenta fantásticamente la velocidad de los objetos que se acercan. Así, un cuerpo incidente de un tamaño dado tendería a producir un cráter mayor en un satélite cercano a Júpiter que el que produciría en otro más alejado. El incremento de la velocidad es despreciable para Calixto, pero viene a ser dos para Ganimedes, tres para Europa y cinco para Io.

Cuando se representa el tamaño de los cráteres en función de la distribución de frecuencia de impacto para los satélites galileanos y para cuerpos representativos del Sistema Solar interno, se obtienen curvas para Calixto y Ganimedes que son muy parecidas a las de la Luna terrestre y Marte [*véase la ilustración de la página precedente*]. En contraste, Europa (basándose sólo en tres probables cráteres) y la Tierra tienen menor número de cráteres de un tamaño dado que los otros cuerpos, entre 1/10 y 1/50 del de aquellos. Aunque en Europa quizás haya muchos pequeños cráteres de diámetro inferior a 15 kilómetros, carecemos de pruebas directas de ello. Como hemos visto, Io no tiene en absoluto cráteres, al menos hasta el límite de resolución de las imágenes Voyager. El bajo valor del número de cráteres en la Tierra indudablemente refleja la dificultad de identificar pequeños cráteres y los efectos de enmascaramiento de los procesos geológicos.

Se pueden establecer las siguientes conclusiones generales. Los terrenos fuertemente salpicados de cráteres de Ganimedes y Calixto, comparables a las tierras altas densamente pobladas de cráteres de la Luna, Marte y Mercurio, deben datar del período de bombardeo torrencial de hace cuatro mil millones de años. Evidentemente, el terreno estriado en Ganimedes comenzó a formarse antes de que terminara el temprano período de intenso bombardeo.

Se puede deducir el límite inferior de la velocidad reciente de formación de cráteres de la manera que sigue. Primero, la superficie de Europa ciertamente no puede anteceder a unos cuatro mil millones de años; de lo contrario, mantendría las huellas de los grandes cráteres antiguos formados durante el período temprano de bombardeo torrencial. Tres cráteres identificados en cuatro mil millones de años establecen un límite inferior para el flujo de objetos capaces de producir cráteres en la vecindad de Europa, límite que es aproximadamente una décima parte del flujo en la vecindad de la Luna terrestre. Esto concuerda con la estima de Shoemaker del flujo en la vecindad de Europa, y podría indicar que la edad de la superficie de Europa ronda varios miles de millones de años. En el otro extremo del margen de posibilidades, la superficie de Europa no puede ser más reciente de unos 100 millones de años; de lo contrario, el flujo en la vecindad de Europa tendría que ser mucho mayor que en la vecindad de la Luna. En cuanto a Io, independientemente de las estimas de flujo, la total ausencia de cráteres simplemente significa que tiene la más joven y dinámica superficie hasta ahora observada en el sistema solar.

Ilusiones geométricas

En psicología son clásicas estas figuras, donde las líneas se ven diferentes de como realmente son. Tales efectos ilusorios parecen guardar relación con la percepción del tamaño de los objetos del mundo tridimensional

Barbara Gillam

Las ilusiones geométricas son figuras compuestas por líneas cuya longitud, orientación, dirección o curvatura se perciben erróneamente. Por ejemplo, en ciertas figuras dos líneas igual de largas parecen tener longitudes distintas. ¿A qué es debido? En la visión tridimensional normal es posible que dos líneas de extensión desigual, situadas a distintas distancias del observador, proyecten sobre la retina imágenes del mismo tamaño; pero, incluso entonces, las líneas son percibidas como distintas, porque el sistema visual consigue reconocer que se encuentran a distinta distancia. Las líneas parecen tener la verdadera longitud que realmente poseen en el mundo tridimensional, debido a que el mecanismo de percepción llamado constancia dimensional consigue compensar las diferencias de distancia, haciendo aparecer más grande la línea más alejada, y más pequeña, la más cercana. Se ha conjeturado que el sistema de constancia dimensional podría ser responsable de ilusiones geométricas. Más concretamente, si el sistema visual procesara las líneas de ciertas figuras geométricas como si fuesen líneas situadas a distintas distancias, el resultado sería una ilusión.

Aunque atractiva, esta explicación de las ilusiones geométricas no es correcta, porque casi todas las figuras que suscitan ilusiones carecen de profundidad, tanto real como aparente. No obstante, varios investigadores, entre los que me cuento, somos de la opinión de que algunos procesos que participan en la exacta percepción de los objetos del mundo tridimensional podrían dar origen a ilusiones en la percepción de imágenes bidimensionales. Mis propios hallazgos indican que el origen de las ilusiones geométricas debe buscarse, no en la apariencia de profundidad de la figura, sino en ciertas claves relativas a la escala y tamaño de objetos del mundo visible; entre ellas, el escorzo y la perspectiva lineal. Posteriormente examinaré con algún mayor detalle la diferencia entre perspectiva y profundidad aparente.

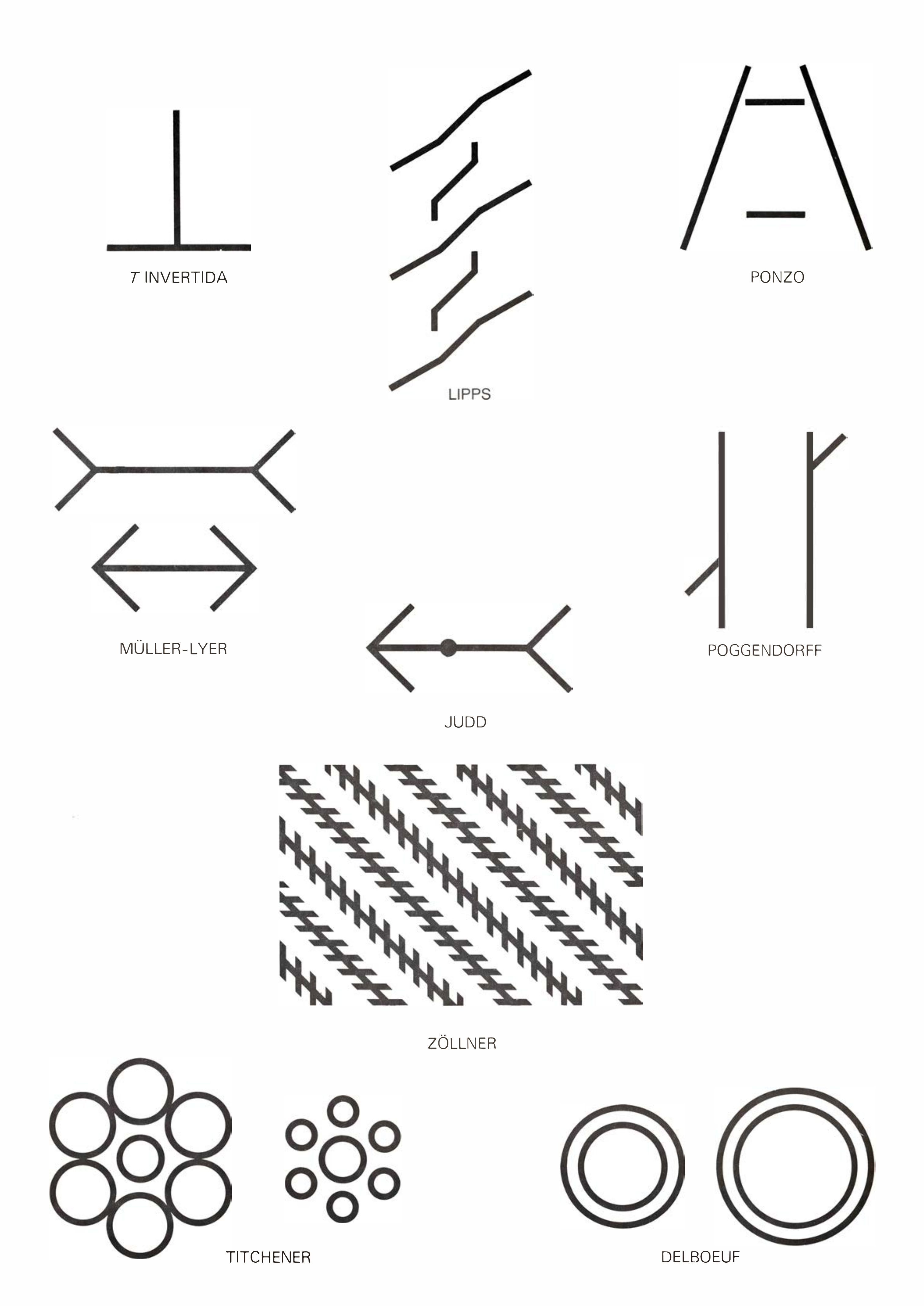
La casi totalidad de las más de 200 ilusiones registradas por los especialistas se descubrieron durante la segunda mitad del siglo pasado. En la ilustración de la página adyacente podemos ver algunas de las más conocidas. En la figura de Müller-Lyer hay dos segmentos, en cuyos extremos se han trazado puntas de flecha, hacia dentro en uno y hacia afuera en el otro. En realidad, ambos segmentos son igual de largos. En la figura de Ponzo también las líneas horizontales son de igual longitud. Tanto en la figura de Zöllner como en la de Lipps las líneas oblicuas son paralelas. En la figura de Titchener, los dos círculos interiores tienen el mismo tamaño. En la figura de Delboeuf, el círculo exterior de la parte izquierda tiene el mismo tamaño que el interior del lado derecho. En la figura de Poggendorff, las líneas oblicuas son en realidad colineales. En la *T* invertida, los trazos horizontal y vertical son igual de largos. Y en la figura de Judd, el punto está situado exactamente en el centro de la recta horizontal.

Tradicionalmente se ha considerado que cada una de estas figuras ilusorias estaba formada por dos componentes: una componente "inductiva", causante de la distorsión, y una componente "de prueba", cuya percepción queda distorsionada. Por ejemplo, en la figura de Müller-Lyer las puntas de flecha forman la componente inductora, y las líneas horizontales, la componente de prueba; en la figura de Poggendorff, las líneas paralelas son la componente inductora, y las oblicuas, las de prueba. Sin embargo, esta división muy bien puede simplificar excesivamente lo que realmente sucede. La distorsión sufrida en la percepción de la componente de prueba no es a veces sino una manifestación –la más evidente– entre una multitud de percepciones erróneas; entre ellas, la de la componente inductora.

A lo largo de los 100 años en que se han venido estudiando las ilusiones geométricas se ha ido proponiendo una serie de explicaciones, de muy variado carácter. Las más convincentes coinciden todas en tres notas fundamentales. Ante todo, las ilusiones no son conceptuales, sino de percepción; es decir, el conocimiento de que un determinado efecto es ilusorio no debilita la fuerza de la ilusión, mientras que casi todas ellas se reducen drásticamente al ser observadas reiteradamente durante un breve período de tiempo. Segundo, las ilusiones no se originan en la retina, pues emergen con fuerza casi plena cuando se le presenta a un solo ojo la componente inductiva, y al otro, sólo la componente de prueba. Por tanto, el origen de las ilusiones debe encontrarse en algún lugar del sistema visual, más allá del núcleo lateral geniculado del cerebro, lugar donde por primera vez se conjugan las señales nerviosas procedentes de los ojos. Tercero, las ilusiones no son consecuencia del movimiento de los ojos: hay experimentos que muestran que las ilusiones, por lo general en toda su magnitud, se presentan asimismo cuando la figura se expone tan brevemente que el ojo carece de tiempo para explorarla, y lo mismo cuando la imagen retiniana de la figura es estabilizada artificialmente mediante un dispositivo especial encargado de mantener fija la imagen en la retina aun cuando el ojo se mueva.

Las explicaciones de las ilusiones pueden agruparse en cuatro categorías prin-

NUEVE ILUSIONES GEOMETRICAS, que se muestran en la página contigua. El segmento horizontal y el vertical de la *T* invertida son igual de largos. Las líneas oblicuas intermedias de la figura de Lipps son paralelas. Tanto en la figura de Ponzo como en la de Müller-Lyer, las líneas horizontales son igual de largas. En la figura de Judd, el punto ocupa el centro del segmento horizontal. En la figura de Poggendorff las rectas oblicuas están alineadas. Las rectas oblicuas de la figura de Zöllner son paralelas. Las dos circunferencias interiores de la figura de Titchener son de igual tamaño. En la figura de Delboeuf, la circunferencia exterior del lado izquierdo es de igual tamaño que la interior del derecho.


T INVERTIDA
LIPPS
PONZO
MÜLLER-LYER
JUDD
POGGENDORFF
ZÖLLNER
TITCHENER
DELBOEUF

cipales: teorías de clasificación, teorías de actividad, teorías fisiológicas y teorías funcionales. Estas teorías no se excluyen mutuamente; en ciertos casos no hacen sino destacar distintos aspectos de un mismo proceso. De estas categorías, la menos ambiciosa es la primera. Las teorías de clasificación tratan de caracterizar propiedades comunes a diversas figuras ilusorias. Lo que tales figuras tengan en común hará pensar en un proceso perceptivo subyacente. Las teorías de clasificación no se dirigen ni al mecanismo ni a la función de estos procesos. Cuando las ilusiones se caracterizan por exagerar la percepción del grado en que la componente de prueba difiere de la componente inductiva respecto de alguna cualidad predominante, como el tamaño o la orientación, se dice que la ilusión es de contraste. La ilusión de Titchener es un ejemplo de contraste de tamaño, mientras que la ilusión de Zöllner estaría encuadrada en contraste de orientación, y, más concretamente, contraste de ángulos.

Las ilusiones se consideran causadas por asimilación o confusión cuando quedan caracterizadas por subestimación perceptual del grado en que la componente de prueba se diferencia de la componente inductora. Dicho de otra forma, el elemento de prueba se distorsiona en el mismo sentido que una cualidad predominante del elemento inductor. La ilusión de Müller-Lyer es un ejemplo de asimilación de tamaño o posición, porque las líneas horizontales se ven alargadas en la dirección de la punta de flecha. La ilusión de Lipps constituye un ejemplo de asimilación de orientaciones, porque la orientación de cada una de las líneas paralelas se percibe desplazada hacia la orientación de sus vecinas. En una misma ilusión pueden manifestarse elementos de contraste y de asimilación. En la ilusión de Ponzo, la longitud aparente de la línea superior queda deformada hacia el contorno (asimilación), mientras que la longitud aparente de la línea inferior se distorsiona como si se encogiera y apartara de él (contraste).

Este esquema de clasificación ha servido para poco más que poner algo de orden en el gran número de ilusiones conocidas. No obstante, sí hace pensar que estos procesos podrían tener elementos comunes con los de percepción del brillo, percepción que acusa también los efectos de contraste y asimilación con respecto al brillo de regiones vecinas en una misma escena. Aparte de que el esquema de clasificación adolece de falta de fuerza para explicar las causas de las ilusiones, tampoco resulta aplicable en algunos casos importantes, como la figura de Poggendorff o la *T* invertida. Además, las líneas que en la figura de Lipps resultan asimiladas, y las que en la figura de Zöllner exhiben contraste, resultan encontrarse en un mismo margen de orientación.

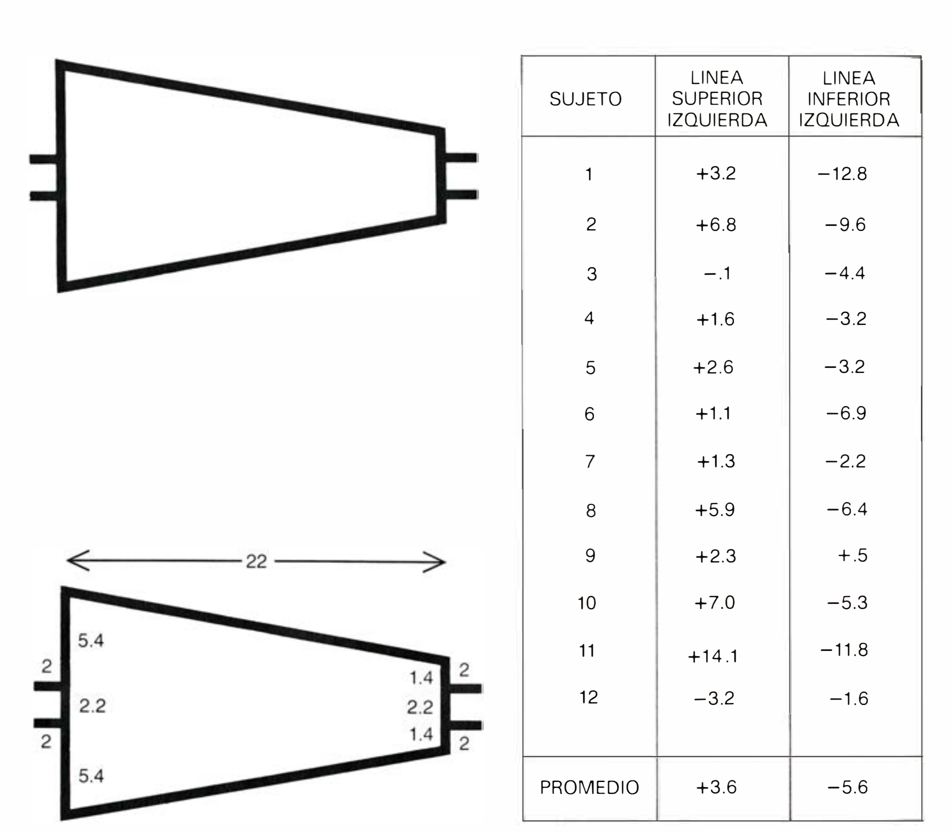
Las teorías de actividad y la teoría fisiológica pretenden ir más allá de la mera clasificación, y afirman que las ilusiones son efectos secundarios fortuitos o errores de la percepción visual normal. Las teorías de actividad atribuyen las ilusiones a respuestas del individuo ante ciertos estímulos, que manifiesta o se dispone a manifestar. La teoría de alerta-eferente, propuesta por Leon Festinger, de la New York School for Social Research, sostiene que se producen ilusiones debido a la forma en que el ojo

RESPUESTA VISUAL A ESTIMULOS RECTILINEOS. Se muestra en forma muy esquemática para dos segmentos cuya orientación difiere en unos 12 grados. Bajo cada uno de ellos se ha trazado la distribución de células detectoras de orientación, situadas en la corteza visual, que han sido activadas por el segmento. La línea de color muestra lo que podría suceder si los segmentos se presentaran simultáneamente al sistema visual. La suma de las distribuciones individuales origina la distribución en color, donde los máximos de actividad han quedado ligeramente desplazados uno del otro. La figura se basa en descubrimientos de Colin Blakemore, Roger H. S. Carpenter y colegas, de Cambridge.

"se pone a punto" para percibir sacudidas y movimientos bruscos. (Como ya he mencionado, las pruebas experimentales obligan a desechar como causantes de las ilusiones los movimientos del ojo, pero tal evidencia no descarta la posibilidad de que los preparativos del sistema visual, encaminados a mover el ojo, sí desempeñen tal papel.) Según dicha teoría las personas tendemos a mirar, no la figura completa, sino ciertas partes de ella, que hacen máximo el número de detalles percibidos con elevada agudeza visual. Festinger y sus colaboradores han descubierto que cuando el observador trata de fijar la vista en los extremos de una figura de Müller-Lyer, en realidad la dirige hacia una zona situada entre las puntas de flecha. Según Festinger, tal proceder tiene por efecto alargar la línea donde las puntas señalan hacia adentro, y acortar la otra, donde las puntas señalan hacia afuera.

Se sabe que el debilitamiento de intensidad de la ilusión, producido por la reiterada observación de la figura ilusoria, va acompañada de mayor precisión en los movimientos del ojo, pero esta correlación no indica que los movimientos oculares sean o no responsables de la ilusión. Otras muchas respuestas motoras, como, por ejemplo, señalar la ilusión con un puntero, podrían suministrar información relativa al error de percepción y, por consiguiente, contribuir a debilitar la ilusión, de la misma forma que lo hacen los movimientos del ojo. Al igual que no sería razonable atribuir la ilusión a la intención de señalar hacia ella con un puntero, sería prematuro, mientras no se disponga de más pruebas, atribuir la ilusión a los preparativos para mover los ojos de manera brusca, o a sacudidas. Además, cuesta trabajo comprender por qué al estar el ojo situado frente a un ángulo o un cruce de líneas habrían de manifestarse imprecisiones de movimiento y acomodación, cuando en el entorno habitual el ojo tiene ocasión frecuente de examinar intersecciones de este tipo.

Las teorías fisiológicas atribuyen las ilusiones a los mecanismos ("hardware") del sistema visual. Muchas de estas teorías se apoyan en el fenómeno de inhibición lateral, que sirve para restringir los estímulos encargados de excitar células de la corteza visual del cerebro. Según estas teorías, la percepción de la orientación de una línea está determinada por los picos (los máximos) de actividad de células especializadas, detectoras de orientación, de la corteza visual, cada una de las cuales es activada por estímulos correspondientes a líneas cuya orientación se encuentre dentro de una determinada gama. Al introducir una



SUJETO	LINEA SUPERIOR IZQUIERDA	LINEA INFERIOR IZQUIERDA
1	+3.2	−12.8
2	+6.8	−9.6
3	−.1	−4.4
4	+1.6	−3.2
5	+2.6	−3.2
6	+1.1	−6.9
7	+1.3	−2.2
8	+5.9	−6.4
9	+2.3	+.5
10	+7.0	−5.3
11	+14.1	−11.8
12	−3.2	−1.6
PROMEDIO	+3.6	−5.6

ESTA VARIANTE DE LA ILUSION DE PONZO (*arriba, a la izquierda*) **está formada por segmentos rectilíneos colineales, pero que parecen estar desplazados. El trazo superior derecho parece demasiado alto para estar alineado con su correspondiente del lado izquierdo, y el trazo inferior derecho, demasiado bajo. La autora pidió a 12 personas que indicasen cuánto debería reajustarse cada trazo del lado izquierdo para dejarlo alineado con su homólogo derecho. A la derecha de las ilustraciones están tabulados los reajustes propuestos por cada observador, en milímetros. En la ilusión de Ponzo ordinaria, de la página 75, la longitud aparente del trazo superior se deforma en el sentido de acercarse al entorno (distorsión llamada asimilación), mientras que la longitud aparente del trazo inferior se deforma en sentido de alejarse del entorno (distorsión de contraste). No obstante, hay en esta ilusión mucho más de qué hablar. La figura superior izquierda muestra que la configuración de Ponzo no sólo influye en la longitud de las líneas, sino también en la alineación de trazos ligados a sus extremos.**

segunda línea de orientación parecida, pero diferente, se origina una distribución ligeramente distinta de la actividad de las células detectoras, actividad que unas veces facilita, y otras inhibe, la generada por la primera línea, produciendo de esta forma desviaciones en la orientación aparente. Con este enfoque, la sinergización, o refuerzo, correspondiente a líneas de orientaciones muy parecidas, provocaría el fenómeno de asimilación en la figura ilusoria, mientras que la inhibición, correspondiente a líneas menos parecidas, provocaría contraste. Esta teoría está respaldada por los experimentos de Colin Blakemore y sus colaboradores, de la Universidad de Cambridge. Para explicar las ilusiones en la percepción del tamaño se ha propuesto una teoría análoga, basada en la existencia de células detectoras del tamaño dentro de la corteza visual.

Las teorías fisiológicas basadas en la inhibición lateral sirven sobre todo para explicar ilusiones donde existe contraste angular y gran número de líneas inductoras, por ejemplo, la ilusión de Zöllner, que suscita asimilación para ángulos extraordinariamente pequeños, y contraste para ángulos un poco mayores. George Wallace y sus colaboradores, de la Universidad de Reading, han descubierto que la ilusión de Zöllner se refuerza introduciendo cambios que puedan causar inhibición lateral, como aumentar el contraste de brillo entre líneas y fondo, o el número de líneas inductoras.

Casi todas las ilusiones donde hay contraste de ángulos persisten cuando la línea de prueba es sustituida por un conjunto de puntos o por un punto móvil. Las teorías fisiológicas basadas en la inhibición lateral no pueden explicar este fenómeno, porque no existe una auténtica línea cuya percepción pueda ser inhibida. Sin embargo, puede que en realidad no haya problema por esta causa, porque se ha realizado muy poco trabajo sobre el tema de si las células detectoras de orientación podrían en efecto responder selectivamente no sólo a líneas sino también a filas de puntos o puntos en movimiento.

Las teorías fisiológicas basadas en inhibición lateral son insuficientes incluso en las más sencillas ilusiones donde in-

tervengan ángulos. Por ejemplo, los ángulos agudos de la figura de Poggendorff, que deberían acusar al máximo el efecto de la inhibición lateral, no contribuyen para nada a la ilusión [*véase la ilustración superior de esta misma página*]. Otra excepción es la ilusión de Müller-Lyer, que ninguna teoría fisiológica ha podido explicar. Al comparar las líneas horizontales de la figura de Müller-Lyer con una línea desprovista de puntas de flecha, se revela que la configuración con las flechas apuntando hacia adentro (ángulos obtusos) contribuye al total de la ilusión de dos a tres veces mas que la otra, donde las flechas apuntan hacia afuera (ángulos agudos). Ninguna organización plausible de las células detectoras de orientación podría originar tan desiguales contribuciones.

Otra teoría fisiológica, la teoría de filtrado selectivo, se basa en la existencia de canales nerviosos en la corteza visual, que al parecer están sintonizados no con estímulos generados por líneas, sino con determinadas frecuencias espaciales de modulación sinusoidal de la luz, siendo la frecuencia espacial el número de modulaciones por unidad de ángulo visual. Ha podido demostrarse matemáticamente que cualquier forma de onda, por compleja que sea, puede ser descompuesta en suma de formas de onda sinusoidales simples. En 1968, Fergus W. Campbell y John G. Robson, de la Universidad de Cambridge, presentaron la teoría de que el sistema visual procesa las formas de onda luminosa descomponiéndolas en sus componentes sinusoidales. Desde entonces se han reunido numerosas pruebas psicofísicas y fisiológicas en favor de dicha propuesta, que actualmente es piedra angular de la teoría visual.

Arthur Ginsberg, de las U.S. Air Force, atribuye las ilusiones a respuestas de canales nerviosos sintonizados a frecuencias espaciales pequeñas. Descompuso las imágenes de ciertas figuras ilusorias en ondas sinusoidales de frecuencias y amplitudes diferentes; eliminó las ondas de alta frecuencia y juntó las de baja frecuencia. Con las componentes de alta frecuencia eliminadas por filtrado, la imagen reconstruida resulta tener propiedades que se corresponden con distorsiones de percepción de la figura original. Por ejemplo, la imagen reconstruida del diseño con las puntas hacia afuera de la ilusión de Müller-Lyer sería tal que el vástago y las flechas se habrían fundido en un segmento difuso más corto, mientras que en la imagen reconstruida de la otra configuración las puntas de flecha y el vástago se confundirían en un brochazo largo y difuso. Las respectivas contracción y dilatación del vástago son, desde luego, la esencia de la ilusión.

Por ingenioso que pueda ser el trabajo de Ginsberg, hay pocas razones para creer que cuando el sistema visual forma juicios dimensionales lo hace filtrando o suprimiendo la información aportada por las formas de onda de alta frecuencia. Sería extraño (aunque no imposible) que el sistema visual juzgase el tamaño aparente de una línea bien diferenciada sirviéndose de las respuestas de canales nerviosos donde la información relativa a la línea estuviese inespecíficamente confundida con la referente a su entorno; concretamente, porque tales juicios inducirían a groseros errores de percepción.

Una de las principales dificultades con que tropiezan tanto las teorías fisiológicas como las de actividad es que no pueden dar cuenta del debilitamiento de las ilusiones. El hecho de que por exhibición reiterada de la figura ilusoria su efecto puede disminuir hasta casi cero durante períodos que pueden durar días, de ninguna manera respalda teorías que postulan mecanismos fisiológicos inmutables. Tan sólo las teorías fun-

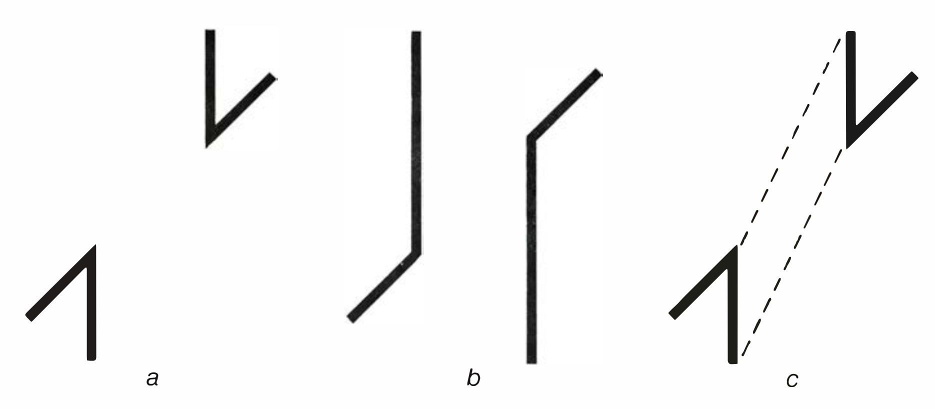
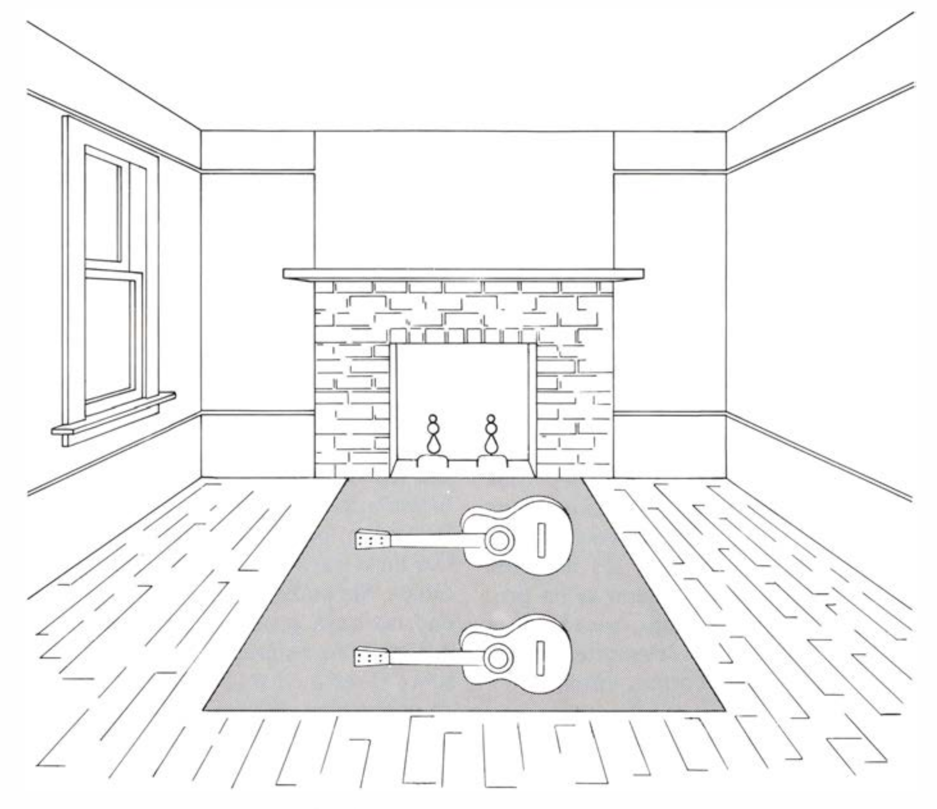


FIGURA DE POGGENDORFF, descompuesta de forma que los efectos ilusorios producidos por las componentes de ángulos agudos (*a*) puedan discriminarse de las correspondientes a los ángulos obtusos (*b*). En *a* la ilusión es nula, o incluso, ligeramente negativa (*a la izquierda*); algunos individuos perciben la línea oblicua derecha como situada algo más baja de lo correspondiente a tramos de una misma recta. En *b* el efecto ilusorio se manifiesta con fuerza (*centro*). Una ilusión en *a* podría deberse a que el sistema visual procesa *a* como si formase parte de *c* (*a la derecha*), donde las líneas de trazos definen un plano que se aleja. En *c* la línea de la derecha se encuentra por debajo de la oblicua izquierda.



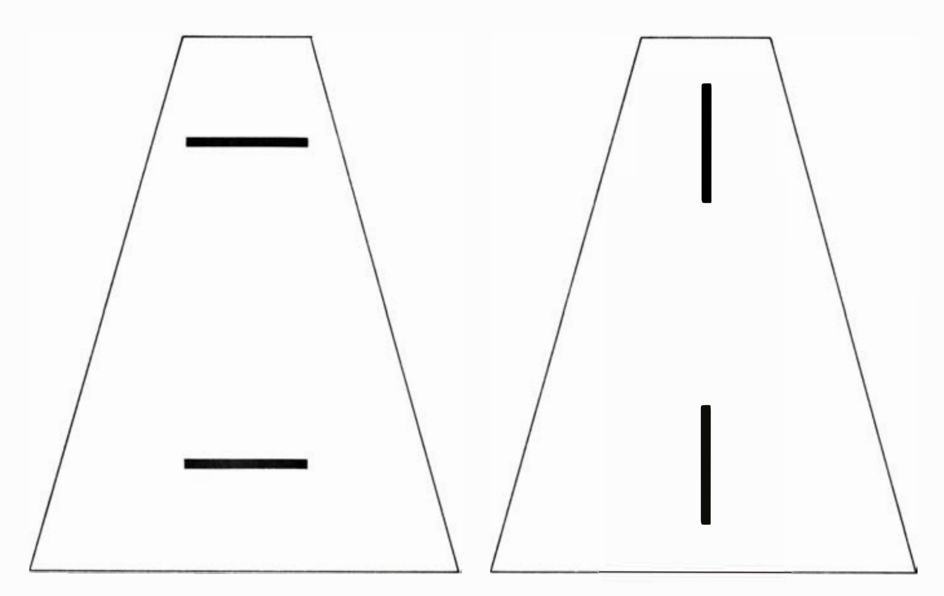
DIBUJO DE UNA SALA; en él reconocemos muchas configuraciones ilusorias que nos son familiares. Una de las guitarras parece mayor que la otra (ilusión de Ponzo); el borde posterior de la alfombra, menor que su dimensión longitudinal (ilusión de la *T* invertida); y el borde anterior más pequeño que la arista de la pared del fondo (ilusión de Müller-Lyer, con la mitad de cada punta de flecha).

cionales, que considera las ilusiones no como errores, sino como procesos esenciales del sistema visual, comienzan a tratar adecuadamente la pérdida de intensidad de las ilusiones.

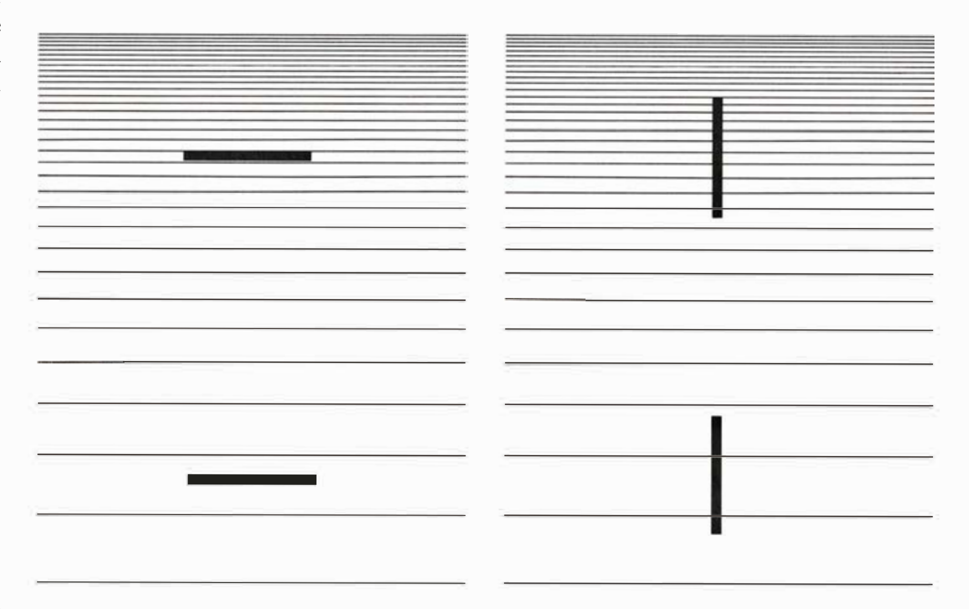
A lo largo del siglo pasado, varios investigadores, entre ellos el psicólogo Armand Thiéry, Richard L. Gregory, de la Universidad de Bristol, y Reinhardt Tausch, de la Universidad de Marburg, pusieron de manifiesto que pueden encontrarse casi todas las figuras ilusorias en representaciones bidimensionales de escenas tridimensionales. En la visión estereoscópica normal, las llamadas distorsiones o ilusiones de las figuras esquemáticas fomentan la percepción exacta del mundo tridimensional y exaltan el realismo de las figuras. Fijémonos en la ilustración inferior de la página precedente. Una de las guitarras parece mayor que la otra (ilusión de Ponzo); el borde posterior de la alfombra parece más corto que su dimensión longitudinal (la ilusión de la *T* invertida), la moldura del techo está demasiado alta para quedar alineada con el zócalo (ilusión de Poggendorff), y el borde anterior de la alfombra parece más corto que la arista de la pared del fondo y el suelo (ilusión de Müller-Lyer con media punta de flecha).

Si bien todas estas percepciones son ilusorias al compararlas con la figura plana o la imagen retiniana, normalmente no las clasificaríamos como ilusiones, pues no son sorprendentes; reflejan características del mundo tridimensional, manifestadas en el plano del dibujo. En la observación tridimensional ordinaria carece por completo de importancia si el zócalo está retinalmente alineado con la moldura, o si el borde de una puerta tiene, en la imagen retiniana, la misma longitud que el rincón del cuarto. Tales hechos, que sólo pertenecen a la figura plana, y dependen sobre todo de la posición del observador, no se registran.

La principal función de la percepción es descifrar la imagen retiniana, transitoria, y conseguir constancia: percepción del mundo exterior expresada mediante sus características estables e intrínsecas. En el dibujo de la habitación, es evidente que las "distorsiones" no son sino ejemplos donde la percepción realiza su función. Sin embargo, si se suprimen los detalles de contexto imprescindibles para obtener impresión tridimensional, las mismas respuestas perceptuales a las mismas configuraciones parecen ahora injustificadas, y son llamadas ilusiones. ¿Significa esto que las ilusiones son únicamente respuestas de constancia (decodificación de la perspectiva) que al operar en un contexto demasiado reducido, por carecer de profundidad aparente, no dan sentido a las ilusiones? Tal punto de vista es atractivo por su sobriedad, mas se enfrenta a una arraigada tradición de la psicología: las respuestas de constancia son provocadas por la impresión de profundidad aparente.



LA PERSPECTIVA LINEAL (compresión de las dimensiones horizontales hacia la parte más alta de un dibujo) puede obtenerse mediante las líneas oblicuas recesivas de la figura de Ponzo. Las líneas de prueba horizontales (*izquierda*) son igual de largas, pero la superior parece más grande, por el efecto de perspectiva. No se produce ilusión para las líneas verticales (*derecha*). En este caso, la percepción del tamaño no es función de la profundidad –ya que ésta aumentaría la longitud de todas las líneas, cualquiera que fuese su orientación–, sino que depende de la perspectiva, que afecta solamente a líneas para las que el fondo suscita gradación de tamaños. (Las ilustraciones son de Jerome Kuhl.)



EL ESCORZO (compresión de las dimensiones verticales hacia la parte superior de un dibujo) se obtiene mediante líneas horizontales paralelas que van aproximándose entre sí conforme se acercan al borde superior del dibujo. Las líneas de prueba verticales (*derecha*) son igual de largas, pero la superior parece más grande por el efecto de escorzo. No se produce ilusión en las líneas de prueba horizontales (*izquierda*). Este resultado confirma que la percepción del tamaño no depende ahora de la profundidad, que alargaría las líneas horizontales y las verticales. La percepción del tamaño parece ser función de la gradación perspectiva del tamaño. El mecanismo de constancia no exige una respuesta de profundidad.

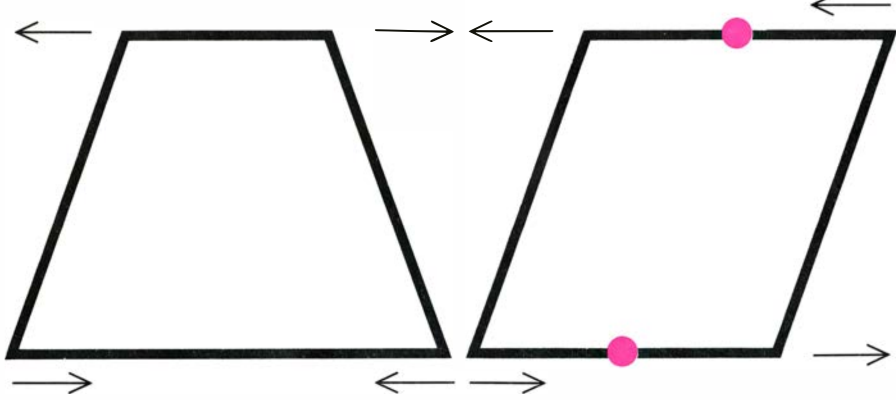
Existe cierto número de motivos para poner en duda que el mecanismo de constancia de escala tenga siempre subyacente una respuesta de profundidad. J. J. Gibson, de la Universidad de Cornell, ha señalado que los gradientes en la compresión aparente de objetos y texturas provocada por la perspectiva aportan información sobre el tamaño y forma de los objetos. Es frecuente que el tamaño de un objeto pueda establecerse por la razón de su anchura a la del fondo. Pensemos en las vías del ferrocarril, convergentes en el infinito. Como el observador sabe que las traviesas distantes son igual de largas que las cercanas, un objeto que abarque los raíles en la lejanía es reconocido como de igual ta-

maño que otro que los abarcase en las proximidades, a pesar de que los tamaños de las imágenes retinianas de ambos objetos sean muy diferentes.

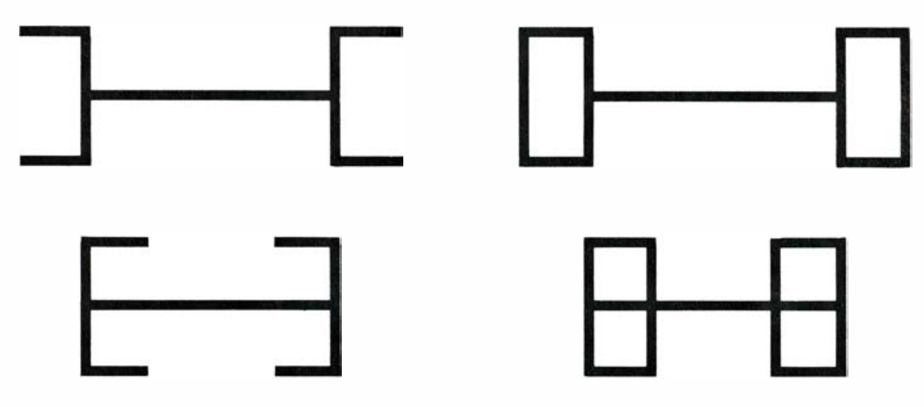
De mis propios trabajos con fondos donde profundidades equivalentes son representadas mediante distintos tipos de perspectiva [*véanse las ilustraciones de la página anterior*], se desprende evidencia concluyente de que existen procesos de constancia dimensional que no son efecto secundario de la percepción de profundidad. En efecto, he conseguido separar la respuesta a la profundidad de la respuesta a la perspectiva. Estudié dos tipos de fondos para las líneas de prueba. En los dibujos del primer tipo, basados en las líneas oblicuas concurrentes de la figura de Ponzo, las dimensiones horizontales son comprimidas más y más hacia el extremo "alejado" (superior) de la figura (deformación llamada "perspectiva lineal"). En los del segundo tipo, basados en líneas paralelas horizontales que van aproximándose entre sí hacia la parte superior del dibujo, las dimensiones verticales son comprimidas cada vez más (deformación llamada "escorzo").

A continuación, situé líneas de ensayo con diversas orientaciones sobre ambos tipos de fondo. Encontré que la dilatación perceptual de las líneas situadas en lo alto con respecto a las situadas en el extremo "cercano", o sea, abajo, solamente se presentaba cuando las líneas de prueba estaban orientadas a lo largo de las dimensiones comprimidas. No se producía ilusión alguna en las líneas de prueba dispuestas según orientaciones para las que el dibujo de fondo no definía graduación dimensional. Dicho de otra forma, una graduación de perspectiva lineal no afecta la longitud aparente de líneas trazadas en el sentido de la dimensión sometida a escorzo. Tal resultado obliga a admitir que existen mecanismos perceptivos que promueven la constancia dimensional basada en la propia escala, y no como respuesta de profundidad asociada a la escala, pues ésta dilataría cualesquiera líneas distantes, sin importar su orientación.

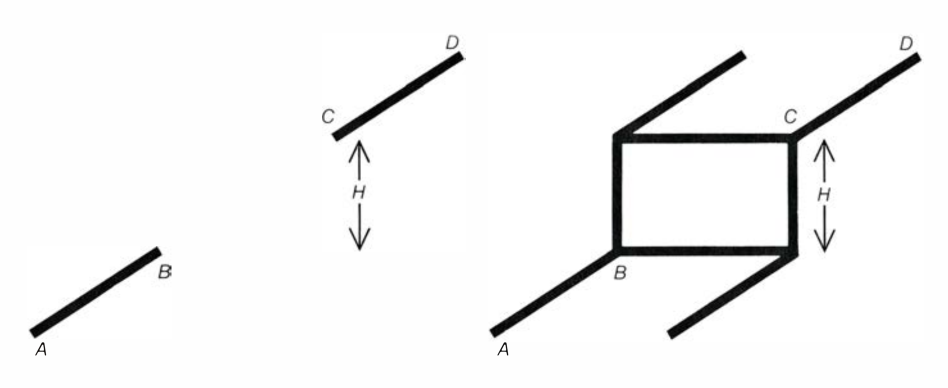
No pretendemos negar que la distancia aparente pueda modificar la percepción del tamaño. Se ha puesto de manifiesto mediante experimentos que, en figuras ambiguas, la inversión del efecto de profundidad lleva consigo modificaciones en la percepción de formas y tamaños. No es sorprendente que la percepción de forma y tamaño pueda quedar afectada por la profundidad aparente, a pesar de ser una respuesta primaria; basta tener en cuenta que la propia percepción de profundidad puede ser influida por el tamaño y forma aparentes. Ejemplo bien conocido es la ilusión lunar. Cuando la luna se encuentra asomando por el horizonte parece mayor que cuando se encuentra en lo alto del cielo. El resultado es que cuando asoma por el horizonte parece estar más próxima. En circunstancias normales, los modos primario y secundario de juzgar tamaños y distancias se refuerzan mutuamente. Empero, mi trabajo muestra que la cantidad de información necesaria para que un dibujo suscite un proceso primario de gradación de tamaños es menor que la necesaria para provocar un proceso de gradación de profundida-



EN LA RETINA, los cuadriláteros que se forman suelen ser proyección de rectángulos inmersos en el espacio tridimensional. Un mecanismo perceptivo que prolongase las líneas horizontales limitadas por ángulos obtusos, y acortase las horizontales terminadas en ángulos agudos, serviría para compensar las deformaciones producidas por la proyección en perspectiva cónica. Las flechas indican las direcciones de la contracción o dilatación perceptual. Los puntos de color señalan el centro de los lados horizontales del cuadrilátero de la derecha. Estos centros parecen demasiado cercanos a los extremos contraídos.



LAS FIGURAS DE MÜLLER-LYER RECTANGULARES (*izquierda*) también provocan una ilusión, aunque mucho más débil que la producida por las figuras ordinarias de Müller-Lyer que se ven en la página 75. La ilusión llega a ser muy fuerte cuando los apéndices se cierran, formando rectángulos (*a la derecha*). Ninguna de las teorías sobre ilusiones geométricas puede dar razón de este hecho.



SEGMENTOS ALINEADOS EN LA RETINA no tienen forzosamente por qué representar rectas alineadas en el espacio tridimensional. Los puntos *B* y *C*, a la izquierda, podrían representar una interrupción en una dimensión horizontal que se aleja continuamente, en cuyo caso la línea *ABCD* estaría contenida en un único plano horizontal en el espacio tridimensional. La otra posibilidad (*derecha*) es que *B* y *C* representen puntos separados horizontal y verticalmente, en cuyo caso *AB* y *BC* yacerían en planos horizontales distintos, y no estarían, por tanto, alineados en el espacio tridimensional. La presencia de líneas verticales paralelas favorece la configuración de la derecha, y por ello, el sistema visual interpreta que las líneas *AB* y *CD* se encuentran situadas a distinta altura.

des. Gregory ha sugerido que este resultado podría deberse al conflicto de percepciones entre la representación de la profundidad de objetos, plasmada en el dibujo, y su carácter auténtico, que es de figura plana.

¿En qué medida puede una teoría funcional, basada no en la profundidad aparente, sino en la perspectiva, explicar ilusiones que, a diferencia de la figura de Ponzo, no exhiban características de perspectiva tan obvias? ¿Cómo explicaría esa teoría las paradojas de Müller-Lyer y de Poggendorff? El parecido de la figura de Müller-Lyer con objetos y escenas que muestran profundidad ha sido reconocida ya desde el siglo XIX. Muy recientemente, Gregory ha hecho notar la semejanza entre las dos configuraciones de Müller-Lyer y una arista de un edificio, una vez vista desde el interior, y otras desde el exterior del mismo. Por otra parte, Tausch y otros investigadores, entre los que me cuento, han descubierto que la ilusión de Müller-Lyer es casi la suma de los efectos que por separado producen las cuatro líneas oblicuas individuales que forman las puntas de flecha sobre la longitud de la línea horizontal, efectos que parecen divorciados de la idea de profundidad. Dos líneas que formen ángulo agudo son perceptivamente más cortas, mientras que si forman ángulo obtuso se perciben más prolongadas de lo que son. Es necesario comprender cómo se distorsionan estas relaciones angulares fundamentales para llegar a entender las ilusiones de Müller-Lyer y otras ilusiones donde haya ángulos que influyan en la longitud percibida de las líneas.

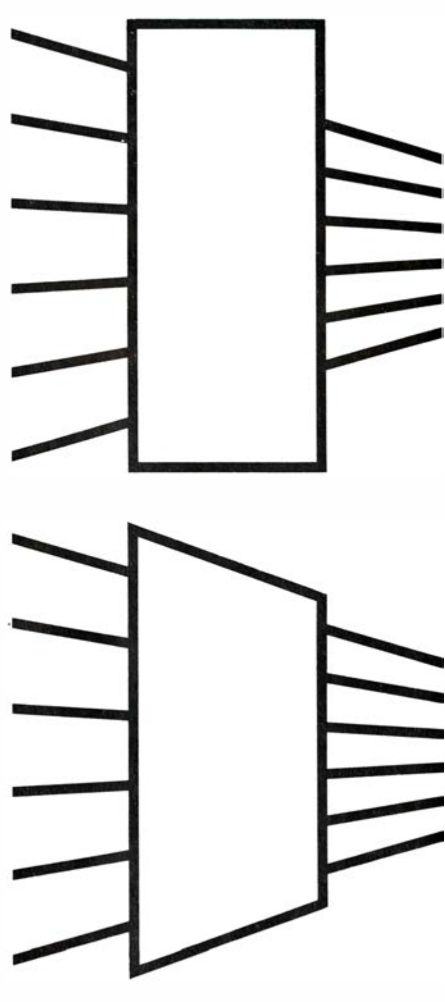
¿Se incrementa la constancia perceptiva por el alargamiento que muestran las líneas limitadas por ángulos obtusos, y el acortamiento de las limitadas por ángulos agudos? En efecto, así ocurre. Una proporción abrumadoramente grande de los ángulos agudos y obtusos que se forman en la retina en el transcurso de la experiencia visual humana son proyección de ángulos rectos. Un mecanismo perceptual que alargase las líneas horizontales limitadas por ángulos obtusos compensaría las diferencias de tamaño creadas por la proyección perspectiva [*véase la ilustración superior de la página precedente*].

La ilustración muestra también que el proceso que acabo de describir facilitaría las constancias de forma y tamaño, tanto si los ángulos obtusos representan las esquinas más distantes y los agudos las más cercanas, como si sucediera lo contrario. Por consiguiente, no debe sorprendernos que la presencia o ausencia de tales distorsiones nada tenga que ver con la percepción de la profundidad. Las partes de la proyección en perspectiva de un rectángulo que forman ángulos obtusos resultan siempre más contraídas, y las partes que formen ángulos agudos, siempre son las más alargadas. El hecho de que los ángulos obtusos tengan mucha mayor contribución a la ilusión de Müller-Lyer sugiere que la gradación de constancia tiende a igualar las longitudes extendiendo las partes más contraídas (limitadas por ángulos obtusos) para igualar en lo posible el resto de la proyección, en lugar de contraer las partes más elongadas (limitadas por ángulos agudos) con el fin de igualarlas con el resto de la proyección.

Si las ilusiones de contracción y elongación sirven para compensar las propiedades de la proyección que no sean propiedades de la escena que fue proyectada, entonces, cuanto mayores sean las distorsiones causadas por la proyección en perspectiva, tanto mayor habrá de ser la fuerza de la ilusión necesaria para corregir la distorsión. Así sucede en la ilusión de Müller-Lyer, que se refuerza conforme los ángulos se apartan más y más del ángulo recto, y conforme las puntas de flecha se van haciendo mayores. (No obstante, el efecto que el tamaño de las flechas tiene sobre la magnitud de la ilusión tiene un límite, del que no puedo dar razón.)

Me propongo ahora estudiar la ilusión de Poggendorff, y analizar su significado desde un enfoque de perspectiva. Las líneas oblicuas de la figura de Poggendorff no parecen estar apreciablemente desalineadas si las miramos solas. ¿Por qué es perturbada en tan alto grado la aparente alineación de las líneas oblicuas por la inserción de las verticales paralelas? Aunque las rectas colineales del mundo tridimensional siempre se proyectan en la retina como rectas colineales, puede ocurrir también que rectas no colineales en el espacio se proyecten como rectas alineadas. Por mi parte, mantengo que las rectas verticales de la ilusión de Poggendorff, y en particular las componentes que forman ángulos obtusos, crean un contexto que sugiere que las líneas oblicuas no representan rectas alineadas en el mundo tridimensional. Así se muestra en la ilustración inferior de la página anterior.

Desde el punto de vista de la perspectiva, las líneas oblicuas representan rectas horizontales que se alejan. Cuando las dos rectas oblicuas están alineadas, como sucede en la ilusión de Poggendorff, los convenios sobre perspectiva dictan que el espacio situado entre ambas en el plano del dibujo podría representar una de las dos disposiciones siguientes. Los puntos *B* y *C* de la ilustración podrían representar una interrupción en una dimensión horizontal que se aleja continuamente, en cuyo caso la línea *ABCD* estaría contenida en un único plano horizontal del espacio tridimensional. La otra posibilidad es que *B* y *C* representen puntos con separación tanto horizontal como vertical; en tal caso *AB* y *BC* yacerían en distintos planos horizontales, y por consiguiente, dichas rectas no podrán estar alineadas en el espacio tridimensional.



ORIENTACION del plano. Dicha orientación modifica enérgicamente el efecto ilusorio en el caso de un haz de segmentos oblicuos con punto de fuga común. Arriba, las líneas son interceptadas por un plano visto frontalmente, y parecen menos alineadas que abajo, donde son interrumpidas por un plano que se aleja hacia el mismo punto de fuga. Esta experiencia aporta prueba de que un contexto significativo en la configuración tridimensional presentada en dos dimensiones por las líneas perturba la percepción de la alineación.

Cuando los detalles situados en el hueco *BC* sean más coherentes con la idea de equidistancia a *B* y *C* que con la idea de diferencia de profundidad de *B* y *C*, al procesar las señales el sistema visual favorece la organización en rectas no alineadas. Las líneas paralelas de la ilusión de Poggendorff son particularmente eficaces a este respecto, porque sitúan a *B* y *C* en un plano vertical, directamente frente al observador. Lo que

en realidad parece suceder es que el contexto perturba la configuración, pasando de una donde *AB*, *BC* y *CD* son vistas cada una de ellas como si poseyeran menor pendiente que la medida sobre el plano del dibujo (pues todas ellas representan líneas horizontales) a otra donde la inclinación aparente de *AB* y *CD* sigue siendo la misma (pues siguen representando líneas horizontales), mientras que la inclinación aparente de *BC* aumenta, pues ya no representa una línea horizontal. El resultado es la ruptura de la alineación percibida. Esta explicación está respaldada por trabajos de Ross H. Day y R. G. Dickinson, ambos de la Monash University, de Australia, quienes hicieron estimar a sujetos la pendiente de *AB*, *BC* y *CD* en la figura de Poggendorff.

Una consecuencia de mi explicación para la figura de Poggendorff es que la ilusión se debilitaría mucho si se cambiase la longitud de las paralelas, de forma que ya no perfilen un plano frontalmente opuesto al observador, sino un plano que se aleja adecuadamente, en cuyo caso los segmentos *AB*, *BC* y *CD* serían colineales en el espacio tridimensional. Tal cambio puede conseguirse haciendo que las líneas oblicuas y los extremos de las paralelas converjan hacia un mismo punto de fuga. Cuando el punto de fuga está definido claramente, la magnitud de la ilusión se reduce a la mitad. Estos experimentos aportan pruebas de que un contexto significativo para la organización tridimensional representada por líneas de un plano puede influir fuertemente en la percepción de alineaciones.

La evaluación de una teoría funcional basada en la decodificación de la perspectiva está dificultada por el hecho de que sabemos muy poco acerca de las respuestas de percepción de la perspectiva, y ello a pesar de que durante 500 años casi todo el arte occidental ha estado basado en la perspectiva. Por otra parte, es poco verosímil que todas las ilusiones sean consecuencia de la decodificación de perspectiva. Por ejemplo, las ilusiones de asimilación de orientaciones no parecen servir a ninguna función de constancia; así sucede con la ilusión de Lipps. No obstante, a diferencia de otras teorías funcionales, mi teoría sí es capaz de explicar de forma general la persistencia de respuestas ilusorias bajo condiciones normales de visión tridimensional, donde las respuestas son reforzadas, pues ahora facilitan la percepción exacta, lo mismo que el debilitamiento de las respuestas ilusorias en condiciones de visión bidimensional, donde las respuestas visuales son inhibidas, pues ahora no cumplen misión alguna.

Queda por investigar la forma en que se produce el debilitamiento, y sobre el papel que en este proceso tienen los movimientos del ojo. Stanley Coren, de la Universidad de British Columbia, y Joan S. Girgus, de la Universidad de Princeton, han propuesto que las ilusiones que se debilitan mucho al ser exhibidas reiteradamente tienen una importante componente de aprendizaje y enjuiciamiento, mientras que las ilusiones que solamente se debilitan levemente, por ejemplo, ilusiones de asimilación como la figura de Delboeuf, son principalmente intrínsecas al sistema visual. Si bien tradicionalmente se ha considerado que las respuestas de constancia estaban basadas en algoritmos aprendidos, no hay razón para que algunas de las respuestas no puedan ser intrínsecas, que evolucionaron precisamente porque generaban constancia. En principio, las teorías que cargan el acento en la función no tienen por qué ser incompatibles con las que resaltan el papel del mecanismo.

Hasta cierto punto, las teorías funcionales basadas en la perspectiva están respaldadas por las respuestas que suscitan las ilusiones geométricas en pueblos y personas cuyo ambiente vital no está dominado por habitaciones, edificios y formas rectangulares. En tales personas las ilusiones tienden a no ser tan fuertes como lo son en gentes de nuestro mismo tipo de cultura. Por otra parte, su respuesta a la *T* invertida no se reduce; al parecer, tal respuesta no es función de la permanencia en ambientes de formas rectangulares, sino del efecto de escorzo producido al observar terrenos distantes. Por sorprendentes que sean estos resultados, no debería concedérseles demasiada importancia en tanto no se establezca claramente que los sujetos que participaron en los experimentos comprendían tan perfectamente como los de nuestra cultura qué se les pedía que hicieran. Empero, si tales resultados demuestran ser perdurables, podrían reforzar las teorías funcionales.

Quizá sean los trabajos efectuados sobre el sentido del tacto los que sirvan de piedra de toque a las teorías funcionales, e incluso, a todas las teorías de ilusión basadas en el procesamiento de señales en el sistema visual. Los experimentos han puesto de manifiesto que cuando los sujetos palpan versiones en relieve de conocidas figuras ilusorias, las ilusiones táctiles que se producen son análogas a las visuales. Empero, en esta etapa del trabajo resulta aventurado establecer identidades entre las percepciones fallidas de carácter táctil y las del sistema visual.

El origen del maíz

El progenitor del maíz moderno es la especie silvestre denominada teosinte. Esta hipótesis, que fue muy controvertida, se apoya en las pruebas aportadas por la mejora vegetal, la arqueología y el folklore

George W. Beadle

El origen y desarrollo de la agricultura en los últimos 10.000 años posibilitó la multiplicación por mil de la población humana. Asimismo, liberó a una fracción importante de la comunidad humana de la tarea de búsqueda de alimentos, permitiéndole participar en otras facetas de la evolución cultural. Resulta bastante sencillo seguir el proceso del origen del trigo, la cebada y el arroz, especies en que se basó el nacimiento de la agricultura del este y oeste asiáticos. En gran parte de Asia se encuentran aún hoy día las especies silvestres de las que derivaron esos cereales, lo que permitió a los genetistas determinar las sucesivas formas de cultivo y la selección y mejora de sus características a que fueron sometidos, primero de manera casual y luego deliberadamente, para que las cosechas resultaran más fáciles y copiosas. Por desgracia, ese hilo conductor no puede seguirse en la planta cultivada de mayor importancia en el hemisferio occidental: el maíz (*Zea mays*). El maíz moderno, con sus mazorcas envueltas por brácteas y sus numerosos granos, no guarda ningún parecido morfológico evidente con plantas del Nuevo Mundo que pudieran reputarse como antepasadas suyas. Ello no obstante, el maíz tuvo que surgir de alguna especie silvestre a través de la mejora selectiva realizada por los antiguos agricultores indios de América central. Si todavía hay o si alguna vez hubo algo así como un maíz silvestre constituye una cuestión científica objeto de fuertes polémicas.

La domesticación del maíz moderno ha avanzado tanto que, en cierto sentido, pudiera decirse que se trata ya de una monstruosidad biológica. Magníficamente adaptada para la producción de grano, la planta no podría sobrevivir en condiciones naturales por no disponer de un mecanismo adecuado para la dispersión de sus semillas. Si se abandonara una mazorca de maíz a sus propias expensas en condiciones favorables para la germinación, originaría gran cantidad de plantitas tan densamente apiñadas que competirían entre sí por el agua y los nutrientes del suelo, de suerte que difícilmente ninguna de ellas llegaría a la madurez reproductora. Sin la intervención del hombre, el maíz moderno acabaría por extinguirse en pocas generaciones. La especie sólo puede sobrevivir si el hombre separa los granos de la mazorca y los siembra.

El origen del maíz trasciende el interés puramente académico. Se trata del cereal más eficiente por lo que respecta a la conversión de energía solar, dióxido de carbono, agua y minerales del suelo en alimento. Ocupa hoy el tercer lugar entre los cultivos destinados al consumo alimentario. Cada año se producen más de 200 millones de toneladas métricas, cifra ligeramente inferior a las cosechas de trigo y de arroz. Aun cuando la fracción de maíz consumida directamente en la alimentación humana es relativamente pequeña, su transformación en carne, leche, huevos y otros productos de origen animal hace de él la planta más importante como fuente alimentaria del mundo occidental. Los maíces híbridos, que representan la mayor parte del maíz que se cultiva hoy en día, tienden a ser susceptibles a las enfermedades; las líneas puras que se cruzan para producir las semillas híbridas son, a su vez, débiles debido a la consanguinidad. Por consiguiente, la correcta identificación de la dotación genética que dio origen al maíz supondría una valiosa fuente de diversidad genética y de vigor híbrido.

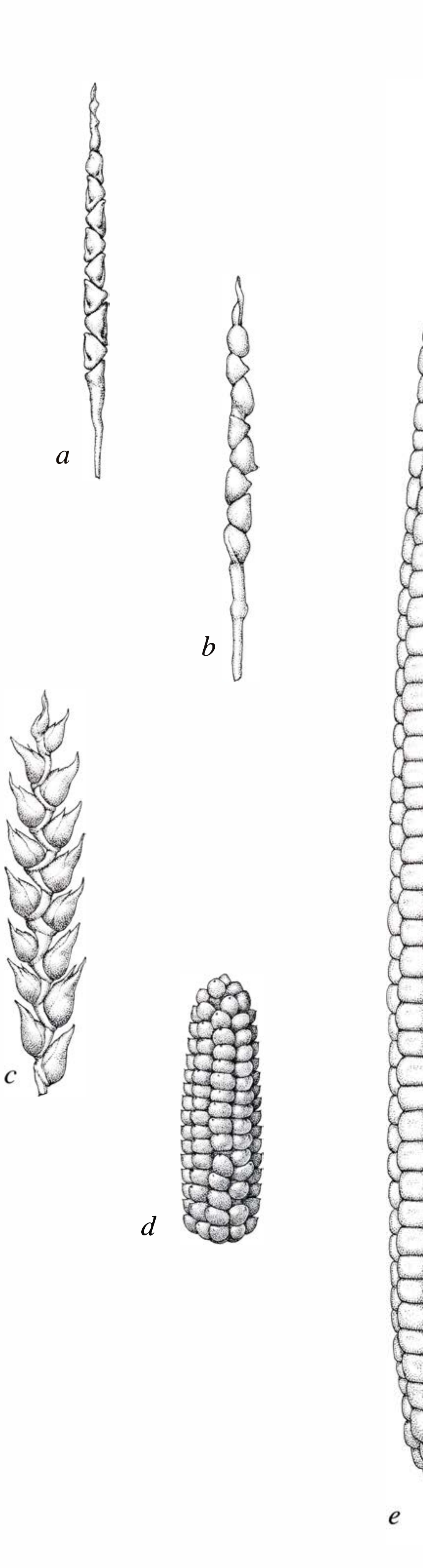
En la actualidad dominan hipótesis enfrentadas en lo referente al origen del maíz. Existe una escuela que sostiene que podemos dar con el maíz silvestre, y que se trata de la que es, y siempre ha sido, la hierba silvestre teosinte. Esta especie crece de manera espontánea en México, Guatemala y Honduras. Esa hipótesis fue muy compartida por los botánicos de los siglos XVIII y XIX, pero perdió predicamento en la primera mitad de nuestro siglo. Para la escuela antagonista, si hubo maíz silvestre, éste quedaría absorbido por posteriores estirpes de maíz que el hombre fue desarrollando. El exponente máximo de esta segunda opinión fue Paul C. Mangelsdorf, profesor emérito de la Universidad de Harvard y actualmente en la Universidad de Carolina del Norte en Chapel Hill, quien publicó un artículo en el número de julio de 1950 de *Scientific American* con el título "The Mystery of Corn". Sin embargo, investigaciones recientes realizadas por el autor de este artículo y sus colegas han aportado nuevas y valiosas pruebas que corroboran la hipótesis del teosinte como progenitor del maíz moderno. Ahora ya parece bastante probable que un teosinte, en un intervalo de tiempo que se remonta hasta hace entre 8000 y 15000 años, fuese el antepasado directo del maíz moderno, y que la transformación en un maíz primitivo se debió a la selección dirigida por el hombre.

El maíz se desconocía totalmente en Europa y Asia antes de la llegada de Colón al Nuevo Mundo, en 1492. Sus hombres descubrieron en la Isla de Cuba extensos campos cultivados con una curiosa planta, nueva para ellos. Más tarde se comprobaría que se cultivaba en todo el hemisferio occidental. El elevado tamaño de las poblaciones aztecas, mayas e incas, con su impresionante nivel cultural, no se hubiera alcanzado sin el maíz, cuya cosecha no sólo les resolvía las exigencias nutritivas inmediatas, sino que, mediante el secado, transporte y almacenamiento del mismo, les servía de alimento durante largos períodos de tiempo. En la lista de tributos que debían pagarse a Montezuma, el postrer emperador azteca, leemos que las veinte provincias de que constaba su imperio aportaban a las despensas reales unos 7.500.000 kilogramos de maíz.

La documentación sobre el maíz más antiguo conocido es de orden arqueológico. Su datación se remonta a unos 7000 años. En pocos milenios, los pue-

blos precolombinos mejoraron la mayoría de las principales variedades del maíz existentes hasta la fecha: maíz colorado, maíz azul, maíz amarillo, maíz forrajero, maíz dulce, maíz dentado, maíz duro, maíz harinoso, maíz tunicado y maíz reventón. En tiempos de Colón se cultivaban de 200 a 300 variedades de la especie. Los antiguos agricultores indios habían trasladado la planta desde su lugar de origen, probablemente el sur de México, y la habían adaptado a una amplia zona de climas de temperaturas moderadas y tropicales, que se extendía desde la desembocadura del río San Lorenzo hasta lo que es hoy Chile central. Aunque los modernos mejoradores han incrementado el rendimiento del maíz a través de nuevas variedades e híbridos adaptados a diferentes regiones y resistentes a enfermedades e insectos, el desarrollo del maíz por los indios sigue siendo el logro más extraordinario que alcanzara el hombre en la mejora de plantas.

Mucho antes de la llegada de los españoles, el teosinte era una planta familiar para las poblaciones de ciertas regiones de México y Guatemala. En 1780 el explorador Francisco Hernández describía el teosinte como una planta parecida al maíz, aunque de semilla en forma triangular. El teosinte se diferencia del maíz en que tiene varios tallos que nacen de la base de la planta, en tanto que el maíz actual muestra normalmente sólo uno. Las semillas del teosinte, que están contenidas en pequeñas cápsulas uniloculares, separadas unas de otras y yuxtapuestas, forman una hilera única de seis a diez granos. Dicha hilera de cápsulas, denominada espiga feme-



DOMESTICACION DEL MAIZ, a partir de la hierba silvestre teosinte. Realizada durante varios milenios por el hombre mediante selección, se ha reconstruido a través de experimentos llevados a cabo por el autor. Todas las muestras están dibujadas a escala natural. La "espiga" de teosinte (*a*) equivale a la mazorca del maíz y consta de una sola hilera de granos encerrados en unas duras cápsulas uniloculares. A la maduración, las cápsulas se rompen y las semillas se diseminan. Los cruces entre teosinte y maíz producen un teosinte modificado (*b*) que podría semejar una antigua forma de transición. Una sola mutación pudo haber dado lugar a una variedad tunicada (*c*), en la cual la dura vaina del fruto se ha transformado en unas blandas glumas en forma de vainas, de las que pueden extraerse fácilmente los granos de maíz. Esta mutación pudo muy bien ser un paso crucial en la domesticación del teosinte. Los cruzamientos entre un teosinte y maíz moderno dan lugar a pequeñas espigas primitivas (*d*), parecidas a las muestras arqueológicas de hace 7000 años encontradas en el suroeste de los Estados Unidos y México. El maíz moderno (*e*) es una monstruosidad biológica creada por la prolongada domesticación. Aunque está bien dotada para producir grano, es incapaz de sobrevivir por sí misma en condiciones naturales, sin el concurso del hombre.

nina, corresponde a la mazorca del maíz. Por otro lado, la planta del maíz produce una sola mazorca, o pocas a lo sumo; no así la del teosinte, en la que típicamente se desarrollan varias espigas. Llegada a la maduración, las cápsulas o frutos individuales de cada espiga del teosinte se desarticulan y diseminan.

La espiga madura del teosinte se asemeja remotamente a la mazorca; ello no fue óbice para que los primeros estudiosos del problema estuvieran ya de acuerdo en que el teosinte fuera el antepasado, más probable, del maíz; se fundaban en que crecía silvestre en las mismas regiones del Nuevo Mundo donde se domesticó por vez primera el maíz. El botánico A. Vinson escribía en 1877: "De acuerdo con la opinión de Darwin, el teosinte es el antepasado del maíz". Pero no compartían tal juicio la mayoría de sus contemporáneos, quienes consideraban tarea imposible que el hombre hubiera conseguido esa transformación.

Más tarde se descubrió lo que parecía ser un candidato más probable: una planta intermedia entre el teosinte y el maíz. Dicha especie presenta una espiga pequeña de granos puntiagudos que sobresalen parcialmente de fundas individuales. Por recordar los dientes del perro, recibió el hombre de maíz de perro o coyote. No tardó mucho en observarse que dicha planta constituía la primera generación de un cruce natural de teosinte con maíz; un cruce al alcance de cualquier hábil experimentador y que da una descendencia fértil. Esta fertilidad de las semillas resultantes de los cruzamientos señalaba que la relación genética entre el teosinte y el maíz era mucho más próxima de lo supuesto.

Ese era el estado del conocimiento en 1928, año en que empecé mi trabajo de graduación con el profesor R. A. Emerson, en la Universidad de Cornell. Hacía poco que Emerson había demostrado que el teosinte era una planta de "días cortos", es decir, que no necesitaba más de 12 o 13 horas de luz solar al día y temperaturas altas. Crece silvestre allí donde se dan esas condiciones. Y había conseguido cultivar el teosinte en la latitud de la Universidad de Cornell, acortando los días de verano de una manera artificial; para lo cual, tapó las plantas con cajas obscuras durante el período crítico de principio de la floración.

En mi calidad de graduando y auxiliar se me asignó el estudio de la citología y genética de los cruces de maíz con teosinte. Con Emerson, confirmé la fertilidad de los híbridos y demostramos que los 10 cromosomas de las células del teosinte mexicano eran altamente compatibles con los 10 cromosomas del maíz. En las plantas híbridas, los cromosomas se emparejaban con normalidad durante la formación de las células sexuales; los nueve cromosomas que pudimos identificar por medio de marcadores formaban parejas e intercambiaban segmentos del mismo modo, en términos esenciales, que lo hacían los de las plantas de maíz. Y llegamos a la siguiente conclusión: citológica y genéticamente, el maíz y el teosinte mexicano podían considerarse una misma especie. La vieja hipótesis de que el teosinte silvestre era el antepasado inmediato del maíz doméstico parecía, pues, totalmente plausible. Dimos por resuelto el problema, en la convicción de que cambios genéticos, relativamente pequeños, convertirían la planta silvestre en doméstica o cultivada, y que era eso lo que a buen seguro había ocurrido.

Guardo buena memoria de las conversaciones con Emerson, cuando decía que sólo se precisarían dos mutaciones para convertir el teosinte en planta útil, desde el punto de vista alimentario. Una que diera lugar a que las espigas, alcanzada la maduración, no se rompiesen; evitando así la diseminación y pérdida de los granos. Otra mutación merced a la cual las cápsulas que contienen semillas resultasen más blandas, de suerte que pudieran desgranarse con facilidad. Tales mutaciones no tenían especial complejidad; algo muy parecido se había registrado en la evolución del trigo, centeno, avena y cebada con relación a sus antepasados silvestres. A pesar de las sólidas pruebas en favor de la hipótesis del teosinte, en las postrimerías de la década de los treinta comenzó a perder adeptos, a raíz de que las investigaciones realizadas en ese campo sugerían entonces que el maíz poseía tantos genes diferentes del teosinte que no podían haberse originado en escasos miles de años de domesticación.

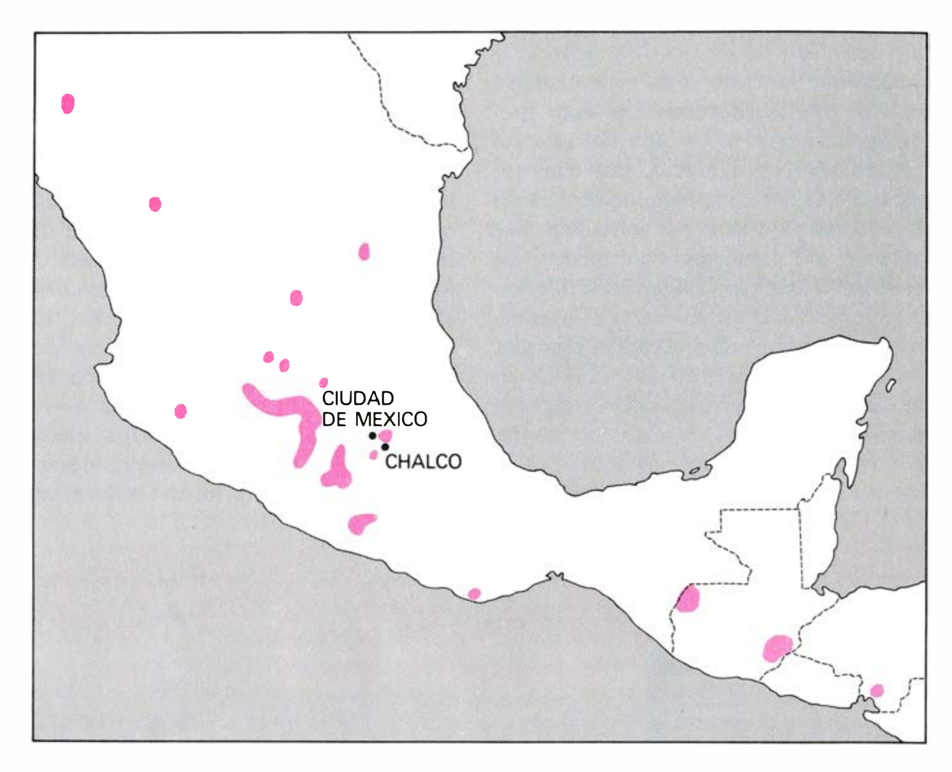
Cuarenta años después de aquellos trabajos con Emerson y tras una vida entregada al estudio de la genética de la mosca del vinagre (*Drosophila melanogaster*) y la genética del moho del pan (*Neurospora crassa*), decidí volver a investigar las relaciones entre el teosinte y el maíz, instado por el descrédito en que la hipótesis había caído. Me jubilé de la

TAMAÑO Y FORMA de una planta de maíz híbrido moderno y de una planta de teosinte. Como puede apreciarse son notablemente diferentes, hecho que justificaba la resistencia a creer que el maíz moderno fuera el producto de la domesticación del teosinte por parte del hombre. El maíz suele tener un tallo único, grueso; el teosinte posee varios tallos delgados. El maíz produce una o dos mazorcas grandes; el teosinte produce numerosas espigas pequeñas, una en cada nudo. Sin embargo, los penachos (flores masculinas) de las dos plantas son similares. El maíz no puede sobrevivir sin la intervención del hombre, en cambio el teosinte está bien dotado para la supervivencia silvestre. (Ilustraciones de T. Prentiss.)

Universidad de Chicago y empecé a trabajar en la determinación del número aproximado de diferencias genéticas existentes entre el maíz y el teosinte. Para ello cultivé a gran escala cruces de teosinte con maíz; poblaciones de híbridos que fueron, con mucho, las más numerosas conseguidas hasta la fecha.

La necesidad para esta clase de trabajo de disponer de gran cantidad de cruces segregantes es una consecuencia de la genética mendeliana. Las leyes de Mendel establecen que, si la diferencia entre los padres es de un solo gen, cada tipo parental aparecerá en las plantas de la segunda generación con una frecuencia estadística de 1 por cada 4. (Nota del traductor: el autor se refiere al caso de que no haya "dominancia" entre tales diferencias genéticas.) La progresión es geométrica. Con una diferencia de dos genes independientes, la frecuencia de reaparición de cada uno de los tipos paternos es de 1 por cada 16, y con 10 genes diferentes la frecuencia se aproximaría a la de uno por millón. Si el teosinte y el maíz divergieran en tan solo 10 genes, yo no podría nunca cultivar y analizar poblaciones de cruces suficientemente numerosas como para producir descendientes con una dotación genética equivalente a la de los padres originarios: maíz y teosinte. Me propuse por tanto cultivar hasta 50.000 plantas de la segunda generación, si fuera necesario. Esa cifra podría darme la oportunidad de encontrar tipos parecidos a los padres originarios en caso de que no hubiera más de seis o siete diferencias génicas importantes. Por ser el teosinte planta de días cortos, que no madura en las latitudes de la zona del maíz de los Estados Unidos, entré en contacto con Edwin J. Wellhausen para que cultivase los híbridos en la estación experimental que el Centro Internacional para la Mejora de Maíz y Trigo tiene en El Batán, cerca de Texcoco, México. El trabajo se hizo en colaboración con Mario Gutiérrez, de la estación El Batán, y Walton C. Galinat, de la Universidad estadounidense de Massachusetts.

Para reducir al mínimo el número de diferencias génicas entre las cepas paternas elegimos una variedad mexicana de maíz, bastante primitiva, llamada chapalote, y la variedad mexicana de teosinte más parecida al maíz, de nombre chalco. Tras varias temporadas de cruzamientos, la frecuencia de reaparición de los tipos parentales correspondientes al maíz y al teosinte en la segunda generación resultó ser de uno por cada 500. Tal frecuencia indicaba que debía haber, aproximadamente, cinco diferencias génicas de importancia, e independientes en su herencia, entre las cepas paternas de maíz y de teosinte: muchas menos de las que se habían sospechado de partida.



EL TEOSINTE TODAVIA CRECE SILVESTRE en zonas de México y Guatemala. No se da en climas templados, por ser planta de "día corto". Necesita solamente de 12 a 13 horas diarias de sol. También requiere altas temperaturas. De acuerdo con la hipótesis del autor, los indios precolombinos conocieron ya de 200 a 300 variedades de maíz adaptadas a una amplia gama de condiciones climáticas.

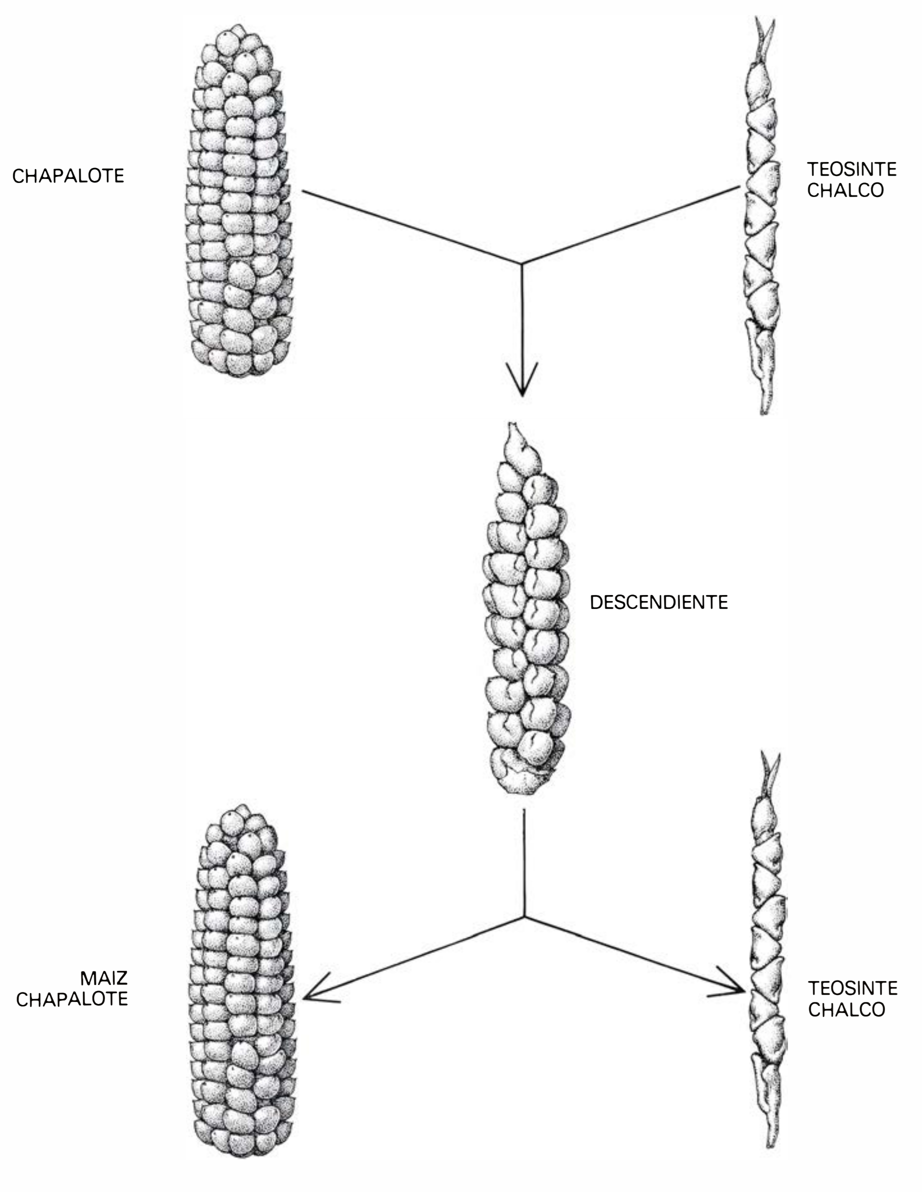
Esos experimentos demostraron con toda claridad que las diferencias genéticas entre el maíz y el teosinte no eran tan grandes como para descartar la hipótesis de una relación ancestral entre ambas especies. Parece, pues, razonable que el hombre precolombino pudo haber seleccionado y conservado los escasos mutantes necesarios para producir un cultivo útil a partir del teosinte, es decir, un maíz primitivo.

Pero Mangelsdorf continuaba en su oposición a la hipótesis del teosinte. En 1939, él y Robert G. Reeves, de la Universidad A. y M. de Texas, postularon que el teosinte no era ascendiente del maíz actual, sino una especie relativamente moderna que procedía del cruce de maíz con una planta del género *Tripsacum*. Su principal razonamiento era el siguiente: el teosinte, por sus caracteres morfológicos, constituía una especie intermedia entre el *Tripsacum* y el maíz. Nunca acepté esa contrahipótesis, y así lo manifesté en 1939. Y ello por varias razones; primero, no se ha observado hasta la fecha que el *maíz* y el *Tripsacum* se crucen entre sí por vía natural espontánea, pese a cultivarse en estrecha vecindad en millones de hectáreas. Los cruces dirigidos por el hombre pueden efectuarse sólo recurriendo a técnicas especiales. Segundo, ningún cromosoma, de los 18 que posee el *Tripsacum*, se empareja normalmente con los 10 del maíz. Y, por último, los cruces de maíz y *Tripsacum* logrados artificialmente por el hombre son completamente androestériles.

La hipótesis del *Tripsacum* ha sido defendida por sus proponentes durante más de un tercio de siglo, y se ha llegado a incluir en muchas obras de genética, botánica, mejora vegetal y demás literatura científica afín. Ello ha inducido que pasara a concebirse, de hipótesis, en un hecho comprobado. Pero dos hallazgos recientes han levantado serias dudas sobre su validez. En primer lugar, y contrariamente a la suposición de Mangelsdorf del origen reciente del teosinte, se acaban de descubrir dos semillas muy parecidas a las del teosinte chalco en un estrato precerámico, no alterado, de hace 7000 años, en un yacimiento arqueológico sito a unos 32 kilómetros al sureste de la ciudad de México. En segundo lugar, Galinat ha demostrado que las duras cápsulas que encierran el fruto del teosinte, y que según Mangelsdorf fueron aportadas por el *Tripsacum* como un bloque de genes ligados, no pudieron surgir así. La razón reside en que las cápsulas de *Tripsacum*, en oposición a las del teosinte, se hallan gobernadas al parecer por genes no estrechamente ligados, por lo que no pudieron transmitirse en bloque.

En respuesta a esas pruebas contra la hipótesis del *Tripsacum*, Mangelsdorf y colaboradores ofrecieron recientemente una propuesta alternativa: el maíz moderno derivaría de un tipo de maíz reventón silvestre, conocido por maíz tunicado o vestido, y el teosinte procedería de ese maíz silvestre por mutación. Sus afirmaciones actuales contradicen las anteriores, que mantenían unas diferencias genéticas entre el maíz y el teosinte tan grandes que aquél no podía proceder de éste. Si el maíz hubiera originado al teosinte, también pudo haberse dado el proceso contrario. Yo diría que lo segundo es mucho más probable, habida cuenta de que el teosinte es una planta silvestre muy adaptada a las condiciones naturales, cosa que no sucede con el maíz.

Si alguna vez hubo un maíz silvestre independiente del teosinte, se extinguió o, a pesar de una búsqueda intensiva del mismo, todavía está por descubrir. Aunque, lógicamente, resulta imposible demostrar que el maíz no existió nunca en forma silvestre, carecemos de pruebas concluyentes que avalen el hecho. En efecto, muchos factores ecológicos hubieran favorecido la supervivencia del teosinte sobre el hipotético maíz silvestre. Por ejemplo, la quebradiza espiga del teosinte es un mecanismo de dispersión de sus semillas mucho más eficaz que lo hubiera sido el carozo de la mazorca del maíz silvestre. El teosinte posee otras muchas características que aumentan sus posibilidades de supervivencia en la naturaleza como planta silvestre. Sus semillas tienen un mecanismo de vida latente que asegura su germinación en cuanto se dan las condiciones favorables para la supervivencia. La dura cápsula proporciona a las semillas del teosinte una gran protección, si la comparamos con los desnudos, o mal protegidos, granos del maíz. Experimentos bien controlados han demostrado que varias especies de roedores, pájaros, insectos y otras especies animales prefieren los granos del maíz a los del teosinte.

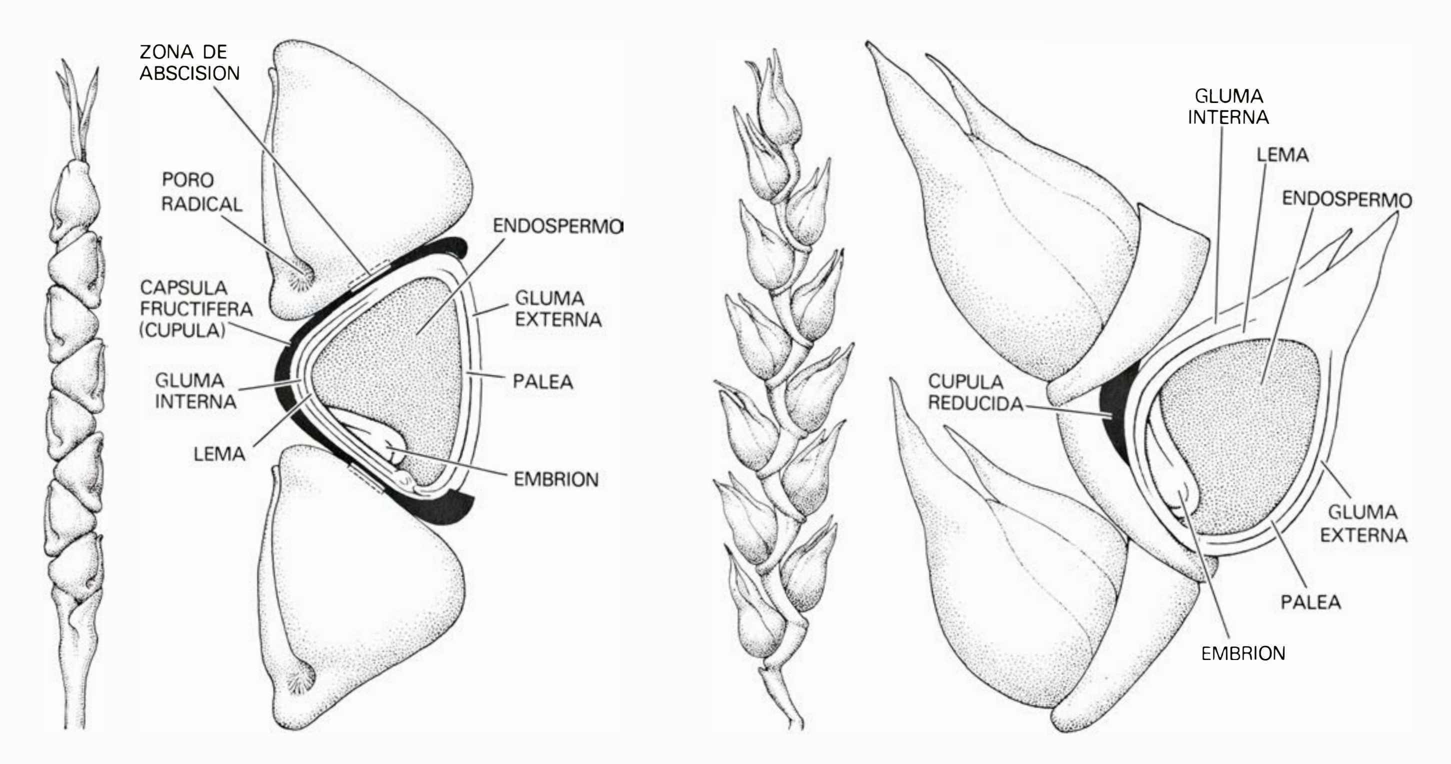


EXPERIMENTOS DE SELECCION, a gran escala, realizados por el autor con objeto de determinar el número aproximado de diferencias génicas entre el teosinte y el maíz. De acuerdo con las leyes de Mendel, si dos variedades relacionadas genéticamente se cruzan, en la segunda generación aparecen formas similares a las de los progenitores del cruzamiento en proporción inversa al número de diferencias génicas entre dichos progenitores. Con objeto de reducir el número de posibles diferencias genéticas, se seleccionaron para el experimento un teosinte parecido al maíz, conocido con el nombre de chalco, y un maíz primitivo denominado chapalote. En 50.000 plantas de la segunda generación de estos cruzamientos, alrededor de una entre cada 500 se parecía estrechamente a uno de los padres; ello manifestaba que el teosinte chalco y el maíz chapalote diferían en cinco genes fundamentales. Resulta plausible, pues, que el maíz se haya originado a partir del teosinte a lo largo de algunos milenios.

Los colores y formas de las cápsulas del teosinte le proporcionan un excelente mimetismo, en virtud del cual alejan la posibilidad de servir de alimento a los animales. El color varía del gris al marrón o al negro; con frecuencia presentan motas, manchas y rayas. Por esto y por su forma ovalada se confunden entre los guijarros y partículas del suelo donde caen. Y lo que es más, el teosinte responde adaptándose a las condiciones del medio, no así el maíz. En suelo fértil y con humedad suficiente, las plantas se desarrollan mucho y producen abundante semilla; en condiciones adversas, por contra, limitan su crecimiento y producen la cantidad de semilla que les permiten los recursos disponibles. El maíz, sin embargo, da una sola espiga y, en condiciones adversas, ni siquiera produce semillas.

Así pues, de haber existido un maíz silvestre, no parece muy probable que compartiera el mismo hábitat o hábitats superpuestos con el del teosinte. Este, con todos los atributos de una planta silvestre bien dotada para la supervivencia, habría desplazado, en la competencia por los recursos, al maíz silvestre. A la inversa, si el teosinte y el maíz silvestre hubieran ocupado hábitats distintos, resulta difícil aceptar que hubieran mantenido la interfecundidad, que la evidencia demuestra, durante miles de años.

¿Cuáles son las pruebas que aduce la arqueología a favor y en contra de esas dos hipótesis antagónicas sobre el origen del maíz? En la década de los cuarenta se recuperaron muchos zuros de mazorcas a partir de antiguas pilas de desechos en cuevas secas del sudoeste de los Estados Unidos y en el valle de Tehuacán, en México meridional. A través del análisis del carbono-14 se demostró que los zuros mexicanos más antiguos se habían cultivado hacia unos 7000 años. En sucesivas generaciones, los hombres que habitaron aquellas cuevas fueron dejando acumularse en el



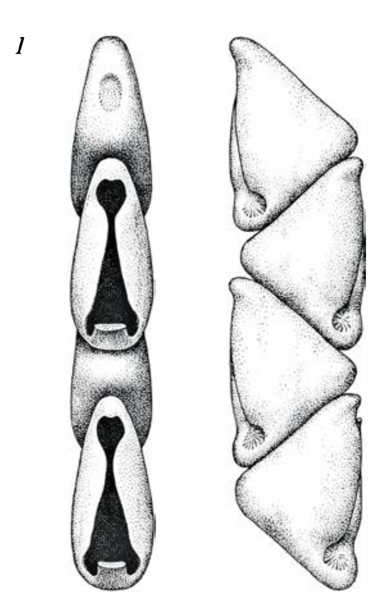
MUTACION "TUNICADA", un paso crucial en la evolución del teosinte a maíz. En el dibujo de la izquierda se muestra la espiga del teosinte; los granos están encerrados en cápsulas duras. En el dibujo de la derecha se muestra el mutante tunicado con sus vainas reducidas a una serie de cúpulas poco profundas y sus granos encerrados en glumas individuales semejando vainas. La forma tunicada tiene menos tendencia a diseminarse cuando madura, lo que facilita que sus espigas se desgranen haciendo de ella una planta alimenticia mucho más útil. Por consiguiente, resulta lógico pensar que una simple mutación fuera capaz de transformar al teosinte en el progenitor del maíz moderno. Posteriores selecciones pudieron entonces haber reducido el tamaño de las glumas, produciendo granos desnudos, susceptibles de ser desgranados sin que los machacara el hombre.

suelo desechos de trilla, que alcanzaron una profundidad de 1,8 metros. Los zuros de distintos estratos revelan una secuencia nítida de los cambios evolutivos provocados por la domesticación. Los zuros más antiguos miden unos tres centímetros de longitud y tienen unos 50 o 60 granos. Mangelsdorf interpretó esos zuros como restos de maíz silvestre auténtico, apoyándose en su uniformidad y en su pertenencia a estratos más antiguos que aquellos otros donde se hallaron restos de plantas que se sabía habían sido cultivadas.

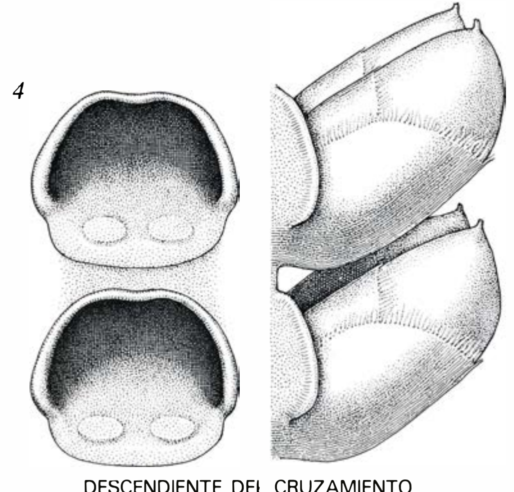
La interpretación alternativa, en mi opinión más plausible, explica que los zuros de mayor antigüedad arqueológica representan estados de transición del teosinte al maíz, a través de un proceso de selección dirigida por el hombre. ¿En qué se apoya esta alternativa? En primer lugar, los zuros antiguos están morfológicamente más próximos al teosinte que al maíz actual. En efecto: algunos de esos primeros zuros son quebradizos y tienen dos hileras de semillas, caracteres ambos típicos del teosinte. En segundo lugar, se obtienen zuros muy similares a los descubiertos por los arqueólogos en la segunda generación y cruces subsiguientes entre teosinte y maíz moderno. En tercer lugar, si los especímenes recuperados en los hallazgos arqueológicos más antiguos fuesen de maíz silvestre y las semillas se hubieran diseminado merced a la fragilidad del zuro, resulta difícil explicar cómo los zuros pudieron soportar, sin fracturarse, la cosecha, el transporte hasta la cueva en que se descubrieron, el desgranado de las espigas y la mezcla con los demás restos del suelo de la caverna. Pero si las mazorcas no fuesen quebradizas, y los zuros en cuestión no parecen haberlo sido, ¿cómo lograron diseminar sus semillas sin la ayuda del hombre?

A pesar de las indiscutibles pruebas en favor del origen del maíz a partir del teosinte, Mangelsdorf y sus seguidores persisten en creer que la dirección de los cambios evolutivos fue de un maíz primitivo al teosinte. La principal prueba que apoya este punto de vista reside en el análisis del polen fósil. El polen del maíz y el del teosinte sólo difieren en el tamaño. El del maíz es significativamente mayor. En 1954, durante las excavaciones efectuadas para la construcción del Museo de Bellas Artes de la ciudad de México, se recuperaron 14 granos de polen en un sondeo a 70 metros de profundidad, cuya antigüedad se calculó en 80000 años. Elso S. Barghoorn, de Harvard, y sus colegas consideraron, basándose en el tamaño, que cinco de estos granos de polen eran de maíz. Barghoorn concluyó que, puesto que el estrato en que se encontró el polen era anterior a la presencia del hombre en el continente americano, los granos de polen grandes debían de proceder de un antiguo maíz silvestre que se había originado independientemente del teosinte. Mangelsdorf escribió más tarde que el polen fósil "mostraba, más allá de toda duda razonable, que el antecesor del maíz cultivado era maíz, no teosinte, ni ninguna de las especies emparentadas con el maíz"

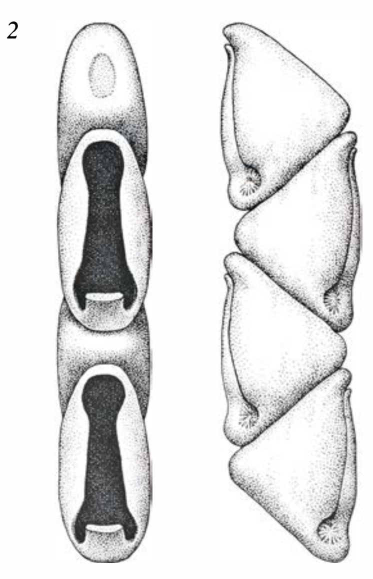
Hay motivos para cuestionar tal interpretación. Existe una estrecha correlación entre el tamaño del polen, la longitud de la espiga y la longitud que debe alcanzar el tubo polínico a través de los pistilos o barbas del maíz para realizar la fecundación del ovario. El maíz primitivo y el teosinte, con sus espigas tenues y polen de pequeño tamaño, no pueden fecundar los ovarios de las grandes espigas de largos pistilos del maíz moderno. Dando por sentado que las espigas del referido maíz silvestre fuesen por lo menos tan pequeñas y tan primitivas como las espigas de hace 7000 años encontradas en el valle de Tehuacán, la correlación entre el tamaño del polen y las pequeñas espigas indica que el polen recobrado en las excavaciones es evidentemente demasiado grande para tratarse de maíz silvestre. Una posible explicación sería que este polen proviniera de un teosinte con cuatro dotaciones cromosómicas (un tetraploide) cuyo polen no hubiera sido reducido todavía por la selección natural al tamaño mínimo efectivo. La aparición de tetraploides es un fenómeno natural.



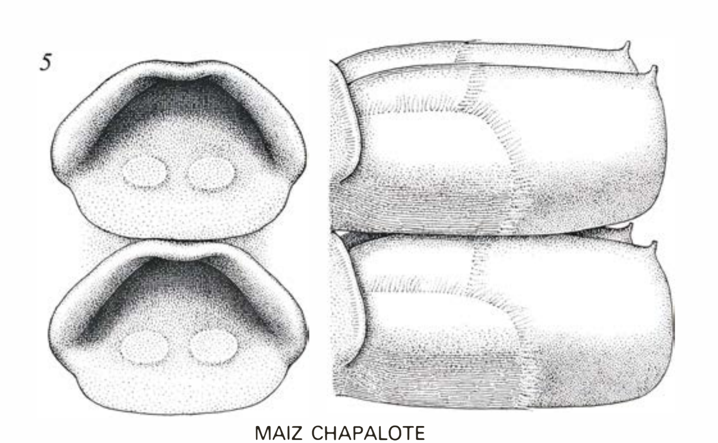
TEOSINTE NOBOGAME



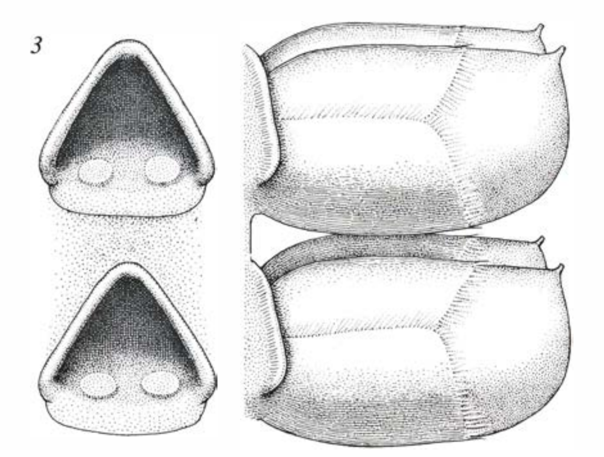
DESCENDIENTE DEL CRUZAMIENTO ENTRE MAIZ Y TEOSINTE



TEOSINTE CHALCO

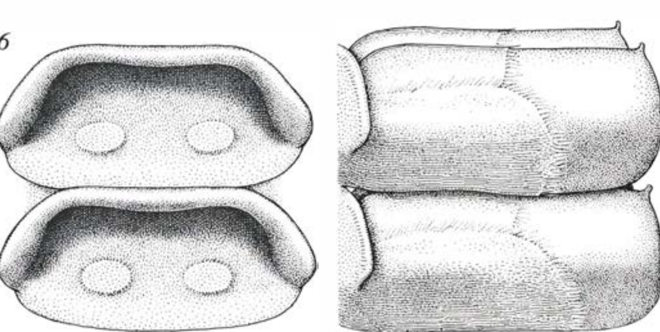


MAIZ CHAPALOTE





MAIZ ARQUEOLOGICO



MAIZ MODERNO

EVOLUCION DE LA CUPULA, desde la cápsula del fruto del teosinte hasta la compacta mazorca del maíz actual. Tal evolución aporta un sólido apoyo al aserto de que el teosinte se halla en el origen del maíz. Se pueden apreciar los progresivos cambios de esta secuencia de su evolución. Primero, los ejes de las cúpulas pasan de una orientación vertical, en el teosinte, a una orientación horizontal en el maíz moderno. Segundo, el espacio entre las cúpulas adyacentes, que corresponde a la parte trasera de la cápsula del fruto del teosinte. se sigue apreciando en el maíz primitivo, pero se va reduciendo gradualmente. En el maíz actual, las cúpulas están próximas, formando un zuro denso y rígido. (Walton C. Galinat, Massachusetts.)

Una segunda posibilidad, más probable, es que el polen en cuestión no sea antiguo, sino resultado de la contaminación de la muestra durante o después de la extracción. Se hubieran requerido extraordinarias precauciones para evitar la contaminación de esas muestras por polen moderno, y es dudoso que tales cuidados se tomasen, dado que el objetivo primitivo de tales excavaciones era el de la construcción. Por estas razones, está claro que la prueba del polen fósil, en el mejor de los casos, no nos lleva a ninguna conclusión.

El argumento que contradice la hipótesis del teosinte como antepasado del maíz es el siguiente: no está claro cómo sus semillas, de cáscara dura, podían servir de alimento al hombre primitivo. Este punto de vista ha persistido a pesar del testimonio aducido por los españoles de que las semillas del teosinte se usaban como alimento, dado que las espigas jóvenes de teosinte pueden comerse antes de que sus cápsulas se endurezcan. En 1939, sin ningún conocimiento previo que acreditase que el teosinte podía comerse, y dada mi afición cuando niño al maíz reventón, a las "palomitas de maíz", se me ocurrió que tal vez las semillas del teosinte también reventarían. Un simple experimento demostró que, en efecto, así ocurría. Los granos reventados, fuera de sus cápsulas, no se distinguían de las palomitas de maíz. Quizás el hombre de la era precerámica preparaba así el teosinte. Los granos colocados en el fuego, en las brasas, en rocas calientes o en arena calentada, revientan muy bien, saltando del fuego o retirándolos de él con un palo o unas tenazas.

He ensayado también la ruptura de semillas secas de teosinte en un primitivo molino de piedra como los usados por los indios precolombinos hace unos 8000 años: el mano y metate. El Museo de Historia Natural de Chicago me prestó amablemente el material arqueológico necesario, de hace unos 8000 años. Llegué a la conclusión de que, en un día, una persona fuerte, instada por el hambre, podría separar mediante un sistema de flotación en agua una cantidad de harina relativamente libre de cáscara de teosinte, suficiente para alimentar una familia pequeña durante un día o algo más. Las semillas maduras de teosinte pueden comerse si se dejan remojar hasta que sus cáscaras ablanden para la masticación. Las cáscaras pueden tragarse o escupirse.

Conozco un mecanismo simple y plausible por el cual el teosinte pudo haber sido transformado por domesticación en una planta alimenticia más útil y en un lógico precursor del maíz. El moderno mutante del maíz llamado maíz vestido o maíz tunicado posee unas hojas blandas que recubren cada grano, las glumas. Mangelsdorf y sus colegas han sugerido que el primitivo maíz arqueológico pudo haber tenido el equivalente genético del carácter tunicado, ya que el maíz arqueológico tiene glumas tunicadas muy apreciables, aunque estas glumas estén reducidas en el maíz moderno a unas escamas en el zuro que no se ven hasta que se extrae el grano. Mangelsdorf empleó maíces tunicados modernos para reconstruir, mediante cruzamientos, los especímenes arqueológicos más antiguos.

Creo que el carácter tunicado desempeñó, en el origen del maíz, un papel más importante que el que se le ha venido reconociendo. Galinat me proporcionó generosamente varias líneas de teosinte con el gen tunicado aportado por el maíz; con estos individuos pude producir muchas líneas adicionales retrocruzándolas con teosinte. Los teosintes modificados, resultantes, tenían muchas de las características que Emerson había considerado propias de los primeros pasos de la transformación del teosinte en maíz. Las cápsulas de los descendientes de cruzamientos entre teosinte y maíz tunicado eran más aplanadas y más blandas, en forma de cáliz, y toman el nombre de cúpulas. Además, las glumas que protegían el grano y lo encerraban en el interior de la cápsula se hicieron mayores y más blandas; la uña podía arañarlas y arrancarlas, dejando el grano al descubierto. También disminuyó la fragilidad de la vaina y la tendencia a diseminar sus semillas al llegar a la maduración.

Bastó, pues, una simple mutación para transformar el teosinte, de cápsulas duras como cáscaras, en una planta cuyos granos desnudos podían desgranarse con facilidad. Si el gen mutante tunicado fue en verdad un primer paso en la transformación del teosinte en maíz, la presencia del carácter tunicado en las especies arqueológicas sería de gran importancia. La plausibilidad de esta hipótesis se ve realzada por el hecho de que el gen responsable del carácter tunicado suele ser dominante. Si en las primeras etapas de cultivo de teosinte hubiese aparecido un mutante tunicado y se hubiera seleccionado para su siembra, no hubiese desaparecido en la siguiente generación, debido a fecundaciones cruzadas, como hubiese ocurrido si el mutante fuese recesivo. Las pequeñas glumas tunicadas del maíz moderno se explican fácilmente. Después del desarrollo de la forma tunicada de teosinte, en subsiguientes fases de domesticación, los antiguos indios mejoradores de plantas desearían reducir el carácter tunicado seleccionando mazorcas con granos desnudos, que pudieran desgranarse sin tener que machacarlas.

Galinat aportó una nueva prueba que corrobora la idea de la transformación del teosinte en maíz. Comparó las cúpulas del maíz arqueológico más antiguo con las del teosinte moderno y con las del maíz moderno. En los teosintes, y a veces en maíces primitivos, las cúpulas son más largas que anchas. En contraste con esto, en los maíces de grano duro del Norte y en sus descendientes modernos, las cúpulas son más anchas que largas. Galinat, al examinar series de maíces arqueológicos procedentes de Nuevo México, de antigüedades comprendidas entre los años 1400 antes de Cristo y 1400 años de nuestra era, observó que estas mazorcas podían disponerse en una línea evolutiva continua desde el teosinte hasta el maíz actual, tomando por base la modificación progresiva de la cúpula. En el teosinte, cada cúpula de una hilera vertical alterna con interespacio que corresponde a la parte trasera de la cápsula adyacente. Este espacio también sufrió un progresivo cambio en la evolución; está presente todavía en el maíz arqueológico pero va reduciéndose progresivamente, de suerte que, en el maíz moderno del área maicera del centro de los Estados Unidos, el zuro se halla apretadamente cubierto de hileras verticales de compactas cúpulas, sin dejar apenas intervalo alguno entre ellas.

La relación entre cápsulas y cúpula parece ser una prueba sólida de que la dirección de cambio fue del teosinte hacia el maíz. ¿Qué otro testimonio podemos ofrecer en favor de la existencia de la cúpula en el zuro del maíz, claramente presente en los especímenes arqueológicos más antiguos? Como ha señalado Galinat, la cúpula evidencia que el maíz procede de un antepasado con cápsula cupulada. Si excluimos el *Tripsacum*, el teosinte es el único candidato. No hay ninguna otra especie viva o extinta que satisfaga tan lógicamente todos los criterios.

Otro conjunto de pruebas que apoyan la hipótesis del teosinte provienen del lenguaje y el folklore. La palabra teosinte deriva del azteca *teocentli*, que significa "la espiga del maíz de Dios". ¿Cómo sabían los aztecas que el teosinte tenía alguna relación con el maíz? ¿Es este un caso de memoria cultural? Además, en muchas partes de México el teosinte es conocido como *madre de maíz*, lo que posiblemente tenga su origen en una denominación nativa más antigua. Atendiendo a la serie de nombres diferentes dados al teosinte por los nativos en su área, el resultado del análisis lin-

güístico podría ser prometedor. H. Garrison Wilkes, de Harvard, encontró en el folklore regional pruebas del papel que desempeñó el teosinte en tiempos prehistóricos. De acuerdo con Wilkes, el explorador C. Lumholz, que viajó mucho por México a finales del siglo pasado, observó que en el área de Nobogame decían que el teosinte cultivado en el mismo campo o cerca del maíz "era bueno para el maíz".

Hay pruebas de que el teosinte cultivado al lado de un campo de maíz incrementa el vigor de éste. La razón es que primitivamente se cultivaba el maíz en pequeñas áreas relativamente aisladas unas de otras, y las semillas se guardaban para la siembra siguiente. La consanguinidad se intensificaba y, después de varias generaciones, el vigor y el rendimiento descendían. Si el teosinte se cultivaba en lugares adyacentes a campos aislados de maíz, o en los mismos campos, se producían cruces entre ellos. La frecuencia de esos cruces dependía de factores tales como la relación espacial y número de polinizaciones. Cuando el teosinte poliniza al maíz, los granos resultantes no se distinguen de los de maíz puro y, así, muchos pueden ser tomados involuntariamente para sembrar. Los individuos de la primera generación muestran un vigor híbrido evidente: nosotros obtuvimos en México descendientes de cruzamientos que tenían de 3,6 a 4,5 metros de altura y que producían hasta 4000 granos por planta.

Aunque las espigas de estos individuos son pequeñas y no deseables, por cuya razón no se toman como semillas para siembras posteriores, estos individuos producirán abundante polen, que se diseminará ampliamente debido a la altura y vigor de las plantas. Los híbridos serán, pues, los padres masculinos en los retrocruzamientos con maíz en la generación siguiente. Algunos de los descendientes de estos retrocruzamientos serán maíces de calidad, mostrando un destacado vigor híbrido y una mayor producción de grano; y en la segunda y subsiguientes generaciones de retrocruzamientos pueden aparecer, en una frecuencia mayor, plantas aún más productivas. De esta forma, el maíz se rejuvenece de acuerdo con el principio empleado en el desarrollo de modernos híbridos de maíz de alto rendimiento.

El rejuvenecimiento de poblaciones de maíz por medio de cruces con teosinte ha quedado demostrado en un experimento muy bien controlado llevado a cabo por Robert J. Lambert y Earl R. Leng, de la Universidad de Illinois en Urbana-Champaign. Cruzaron una línea pura de maíz, bien adaptada y de gran rendimiento con cuatro variedades de teosinte; repitieron retrocruzamientos sucesivos con la línea pura de maíz. Los rendimientos en el segundo y tercer retrocruzamiento se incrementaron en un cien por cien. Estos cruces no son infrecuentes en lugares de México donde el maíz y el teosinte se cultivan juntos. En estas poblaciones los genes de maíz y teosinte se intercambian, regularmente, con una baja frecuencia, debido a lo cual mantienen el vigor híbrido.

Con las prácticas modernas de producción de semilla de maíz, el vigor híbrido es controlado y aumentado al máximo sin la repetida introducción de genes de teosinte. A la larga, sin embargo, hará falta quizás una diversidad genética más amplia, por ejemplo para aumentar la resistencia a enfermedades, y el teosinte podría muy bien suministrarla. En 1973 se encuestó a cuarenta mejoradores y genetistas de maíz sobre la conveniencia de crear un banco de semillas que abarcase una amplia serie de variedades de teosinte y de individuos de la segunda generación, o de generaciones posteriores, de cruces de maíz y teosinte. De los encuestados, 35 consideraron tales semillas como una valiosa fuente de diversidad genética que probablemente sería útil en la década siguiente, cinco consideraron la propuesta interesante pero de dudoso valor científico para la mejora, pero ninguno negó la validez de conservar la semilla. Una encuesta similar sobre la necesidad de preservar el teosinte silvestre hubiese ciertamente logrado una idéntica o aún más positiva respuesta, ya que una gran cantidad de poblaciones silvestres de teosinte de Guatemala y México han sido ya destruidas. Las que quedan están en grave peligro de extinción, ya que el teosinte es muy apetecido por los animales de pastoreo, salvajes o domésticos.

Poco se ha hecho en relación con la conservación del teosinte, aparte de este banco de semillas un tanto fortuito. Ello se debe, en parte, a que el teosinte es una planta de "días cortos", de no fácil cultivo en las áreas templadas del maíz. Se dispone actualmente de una cierta cantidad de teosintes genéticamente adaptados, de fotoperiodismo neutro, que pueden alcanzar gran importancia en programas de mejora más extensos. Sería por tanto una medida inteligente proteger al teosinte, estableciendo reservas en latitudes y altitudes apropiadas, en áreas especiales, como son los parques, yacimientos arqueológicos, jardines botánicos, facultades de agricultura y en las estaciones de mejora oficiales y privadas. Los costos en dinero y esfuerzo serían insignificantes y los beneficios, enormes.

La rabia

Este antiguo azote de la humanidad persiste a través de animales salvajes infectados, principalmente mofetas, murciélagos y zorros. El tratamiento clásico, con la vacuna de Pasteur, puede reemplazarse por una serie de inoculaciones

Martin M. Kaplan e Hilary Koprowski

El 21 de mayo de 1908, el *North American* de Filadelfia publicaba el siguiente titular: "Narcotizado por propia voluntad, víctima de hidrofobia muere tranquilamente". El artículo narraba la dramática historia de un acomodado industrial que había cursado estudios de medicina y sabía que tras la aparición del espasmo laríngeo, uno de los primeros síntomas de la enfermedad, el curso de la misma era inevitable. Dio cuenta al médico que lo atendía de la posible entrada del virus (se había rascado un eccema en el muslo después de haber curado a su perro de una herida por mordedura), arregló sus negocios y pidió opio para tratar las convulsiones que antecederían a la muerte. Pero ya Cornelius Celsus, del primer siglo de nuestra era, había prescrito un tratamiento basado en la inmersión en una piscina, y, de haberse seguido las medidas preventivas de Celsus, que incluían la incisión inmediata de la herida y la cauterización de la misma con un hierro al rojo, la víctima se hubiera salvado.

Sobre la rabia existen referencias anteriores a Celsus. Así, Aristóteles, del siglo IV antes de Cristo, estableció erróneamente que el hombre era el único ser que no contraía la rabia por mordedura de un animal enfermo. En la Ilíada (más de 700 años a. de J.C.) Aquiles llama a Héctor "perro rabioso" y en Mesopotamia, entre las "Leyes de Eshnunna" (1800 años a de J.C.) encontramos: "Si un perro está enfermo y las autoridades se lo han notificado al dueño; si este no lo recoge y el perro muerde a un hombre que muere a consecuencia de ello, el dueño deberá pagar dos tercios de una mina de plata (40 shekels). Si muerde a un esclavo y le ocasiona la muerte, pagará 15 shekels de plata".

En 1546, el médico italiano Girolamo Frascatoro hizo una descripción detallada de la patología del hombre afectado de rabia: "Su incubación [tras la mordedura de un animal rabioso] es tan solapada, lenta y progresiva que la infección raramente se manifiesta antes de los 20 días, en la mayoría de los casos después del día trigésimo y en muchos no lo hace hasta transcurridos cuatro o seis meses. Se han descrito casos en los que los primeros síntomas no aparecieron hasta pasado un año de la mordedura". Una vez que la enfermedad se ha establecido, "el paciente no puede permanecer ni de pie ni tumbado y como un loco se mueve de acá para allá, se araña y siente una sed intolerable. Este es el síntoma más desagradable ya que rehúye el agua y todos los líquidos, y preferiría morir antes que beber o acercarse al agua. Es entonces cuando muerde a otras personas, echa espuma por la boca, tuerce los ojos y finalmente queda exhausto y pacíficamente expira".

El tratamiento moderno se basa en la aplicación, tras la exposición a un animal que se sepa o se sospeche que está rabioso, de una vacuna que consiste en una penosa serie de 14 a 21 inoculaciones. Cada año más de un millón de personas, entre ellas más de 30.000 en los Estados Unidos, tienen que someterse a este procedimiento desagradable. Afortunadamente se está sustituyendo rápidamente en varios países por un tipo de vacuna diferente que Tadeusz, J. Wiktor, nosotros y otros investigadores desarrollamos en el Wistar Institute of Anatomy and Biology. Requiere sólo cuatro o seis inoculaciones, sin que hasta ahora se hayan manifestado efectos colaterales.

El miedo del hombre a la rabia está ampliamente justificado, pues son pocas las personas que han logrado recuperarse una vez aparecidos los primeros síntomas de la enfermedad. El gran autor y dramaturgo austríaco Ferdinand Raimund se pegó un tiro el 25 de agosto de 1836, al día siguiente de haber sido mordido en la mano por un perro que había atacado a otras personas y con el que más tarde acabó un policía. Raimund había sentido un miedo espantoso a la rabia desde su infancia, hasta el extremo de que en cierta ocasión llegó a pensar que había contraído la enfermedad porque un perro lamió el trozo de pan que él estaba comiendo.

Una medida de defensa eficaz contra la rabia requiere que, tras la mordedura, el tratamiento comience lo antes posible. A pesar del éxito de las vacunas y de que más de cien laboratorios están investigando sobre la rabia, aún no se comprende bien cómo y cuándo el virus produce la enfermedad y en qué consiste la acción protectora de las vacunas.

Aunque la rabia puede afectar a una gran variedad de mamíferos, en el pasado el principal reservorio y transmisor de la enfermedad fue, en la mayoría de los países, el perro. Los griegos denominaron a la enfermedad *lyssa* (locura) y los romanos utilizaron la palabra *rabere* (rabiar), de donde deriva el término actual. En la antigüedad la enfermedad se asociaba a la aparición de Sirius (estrella del perro) en los "días del perro" del verano, pues se suponía que los perros estaban entonces particularmente expuestos a la enfermedad.

Durante siglos se ha reconocido que el agente causal de la rabia estaba relacionado con la saliva de los perros enfermos, pero hasta principios del siglo XIX no se pudo demostrar experimentalmente dicha relación (lo que supuso un gran paso en la historia de las enfermedades infecciosas). En 1804 el investigador alemán G. Zinke demostró que la enfermedad podía transmitirse a perros sanos por inoculación de saliva de otros rabiosos, y en 1821 François Magendie y Gilbert Breschet infectaron canes con saliva humana, demostrando que en ambos casos estaba implicado el mismo agente. Sin embargo, hasta hace unos 20 años, cuando comenzaron a utilizarse las técnicas modernas de microscopía electrónica y el cultivo de tejidos, no se pudo caracterizar el agente vírico causal.

Al microscopio electrónico el virus de la rabia recuerda una bala. Su estructura

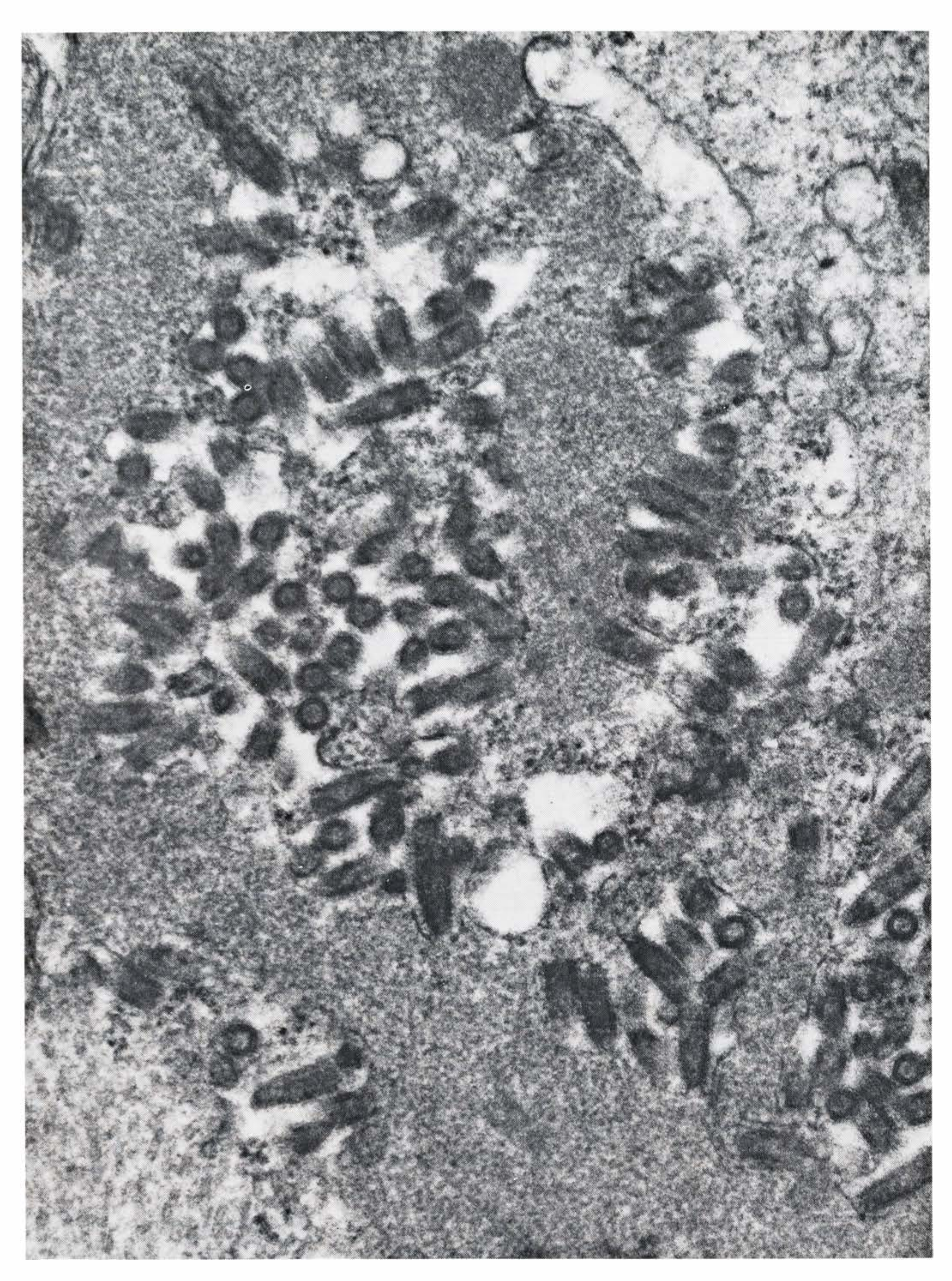


IMAGEN DEL VIRUS RABICO a 116.000 aumentos. La micrografía muestra un corte del tejido cerebral de una rata infectada de rabia. En muchas de las figuras sombreadas se puede observar la característica forma de proyectil. Las formas oscuras circulares son secciones transversas de partículas víricas. La microfotografía electrónica superior fue realizada por Carey Calaway del Center for Disease Control, U.S. Public Health Service.

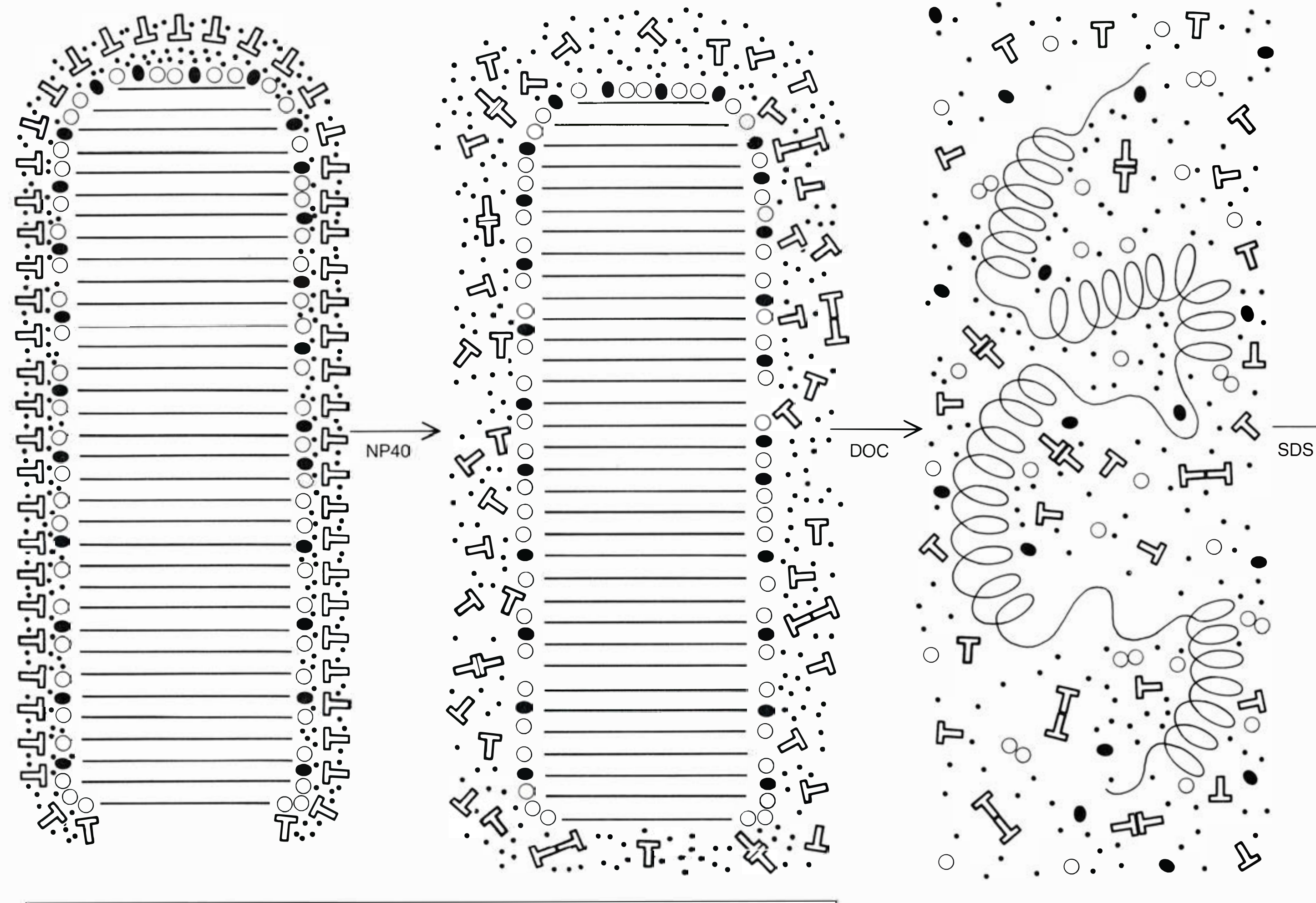
química está constituida por lípidos, proteínas y ácido ribonucleico (ARN). Taxonómicamente se incluye entre los rabdovirus, amplio grupo de agentes víricos que atacan a las plantas, insectos y animales tanto de sangre fría como caliente.

Hasta fechas recientes, debido a la carencia de reactivos apropiados que pudieran identificar los antígenos específicos determinantes de los distintos componentes de los rabdovirus, resultó imposible estudiar el problema de la interrelación y evolución de estos virus. Hoy día, por el contrario, la disponibilidad de anticuerpos monoclonales, y en particular de aquellos que detectan los determinantes antigénicos de la nucleocápside de los virus del grupo de la rabia, posibilita el estudio de las interrelaciones entre varios rabdovirus, lo que permite discernir mejor sobre la evolución a partir de un antecesor común.

Hasta hace muy poco se pensaba que los antígenos representados por las glucoproteínas y las proteínas de la nucleocápside eran similares, si no idénticos, prescindiendo de la cepa de virus rábico de la que se aislaban, en estado relativamente puro. Esta suposición se basaba en que los antígenos presentaban una inmunorreactividad similar frente a sueros de animales inmunizados con el virus rábico.

Al estudiar el virus rábico con la técnica de inmuno-ensayo con anticuerpos monoclonales específicos de la rabia, obtenidos por hibridaciones en cultivos de tejidos, se demostró que cepas de virus rábico hasta entonces consideradas idénticas, o estrechamente relacionadas entre sí, podían ser antigénicamente bastante diferentes. Por ejemplo, dos cepas de virus fijo adaptadas al laboratorio procedentes del virus rábico aislado originalmente por Louis Pasteur, diferían entre sí y ambas del virus de Pasteur desde el punto de vista antigénico; y la cepa Kelev, más atenuada, reaccionó sólo con un hibridoma de los seis ensayados, productores todos de anticuerpos monoclonales. En contraste con las marcadas diferencias antigénicas observadas entre las cepas de virus fijo adaptadas al laboratorio de la rabia, diferentes cepas de virus de "la calle" (aisladas de animales infectados naturalmente en distintos países) mostraron una especificidad antigénica similar cuando se ensayaron con un panel de anticuerpos monoclonales.

Los resultados obtenidos con anticuerpos monoclonales también reflejaban diferencias en la composición antigénica de las glucoproteínas de las cepas de virus rábico. A. Flammand, Wiktor y uno de nosotros (Koprowski) estudiaron las reacciones de los anticuerpos monoclonales frente a la nucleocápside aislada a partir de varias cepas de virus rábico. Contrariamente a lo que se esperaba, descubrieron que las proteínas de la nu-



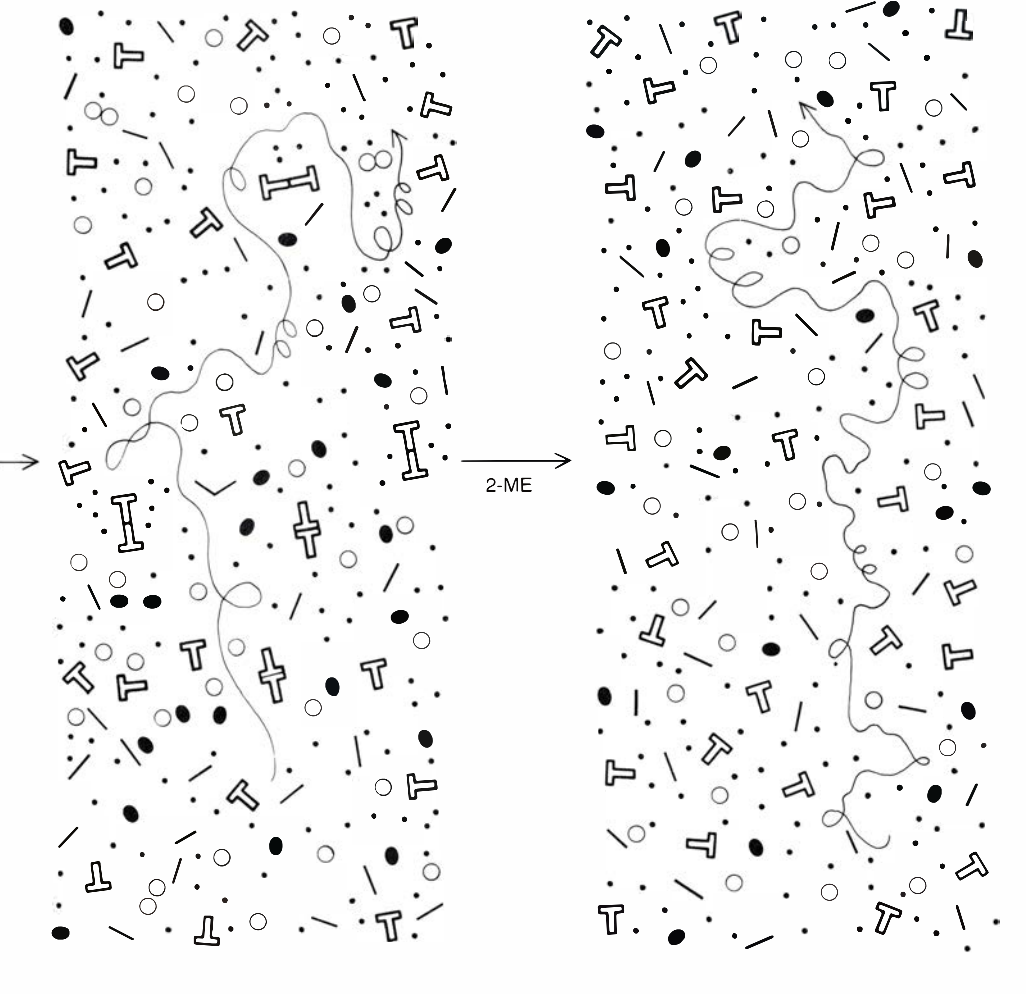
ESTRUCTURA DEL VIRUS, esquematizada comenzando por el extremo izquierdo. El núcleo cilíndrico interno está rodeado por dos envolturas proteicas, aquí numeradas 2 y 3 para diferenciarlas de las glucoproteínas que recubren el virus. Mediante tratamiento con un detergente (NP40) las glucoproteínas se desprenden. El tratamiento

cleocápside pueden clasificarse en diferentes grupos antigénicos. Estos resultados nos han permitido también establecer la existencia de determinantes antigénicos comunes a la nucleocápside del virus rábico y de virus que se conocen como virus rábicos asociados. Dos de estos virus son el Makola y el Lagos Bat, de origen humano y de murciélagos, respectivamente, aislados en Nigeria. En contraste con el virus rábico se ha comprobado que el virus Makola produce en el hombre una enfermedad no letal. Otra cepa africana de virus asociado a los virus de la rabia, la cepa Duvenhage, se aisló de un hombre agonizante en Sudáfrica. Se ha comprobado que las glucoproteínas de esta cepa reaccionan solamente con los dos hibridomas que reaccionan también con la nucleocápside del virus rábico.

Los anticuerpos monoclonales pueden producirse, potencialmente, a un coste relativamente bajo y en grandes cantidades. Estos anticuerpos podrían reemplazar a los sueros animales utilizados para el diagnóstico de la rabia en células infectadas, toda vez que aquellos poseen mayor especificidad de reacción que los anticuerpos de los citados sueros. También podrían llegar a sustituir, en el tratamiento de la enfermedad, a los sueros antirrabia de caballo o de mulo, que causan reacciones locales desfavorables en un 15 a 20 por ciento de las personas inoculadas. Podrían también desplazar a las gammaglobulinas antirrábicas de origen humano, caras y difíciles de obtener, en el tratamiento postinfección de las personas mordidas por animales rabiosos.

Aunque el proceso de la replicación vírica en una célula infectada está perfectamente dilucidado, queda aún por desentrañar qué ocurre inmediatamente antes, una vez que el virus ha sido depositado en el tejido corporal (normalmente en una herida producida por mordedura). Recientemente, Frederick A. Murphy, del Center for Disease Control en Atlanta, ha encontrado la primera prueba de replicación viral en células musculares cercanas a la herida de entrada en hamsters recién nacidos. Esta replicación puede representar un aumento de las partículas víricas capaz de crear una dosis infectante, si bien el inóculo original de partículas víricas en la saliva puede ser suficiente para dicho propósito.

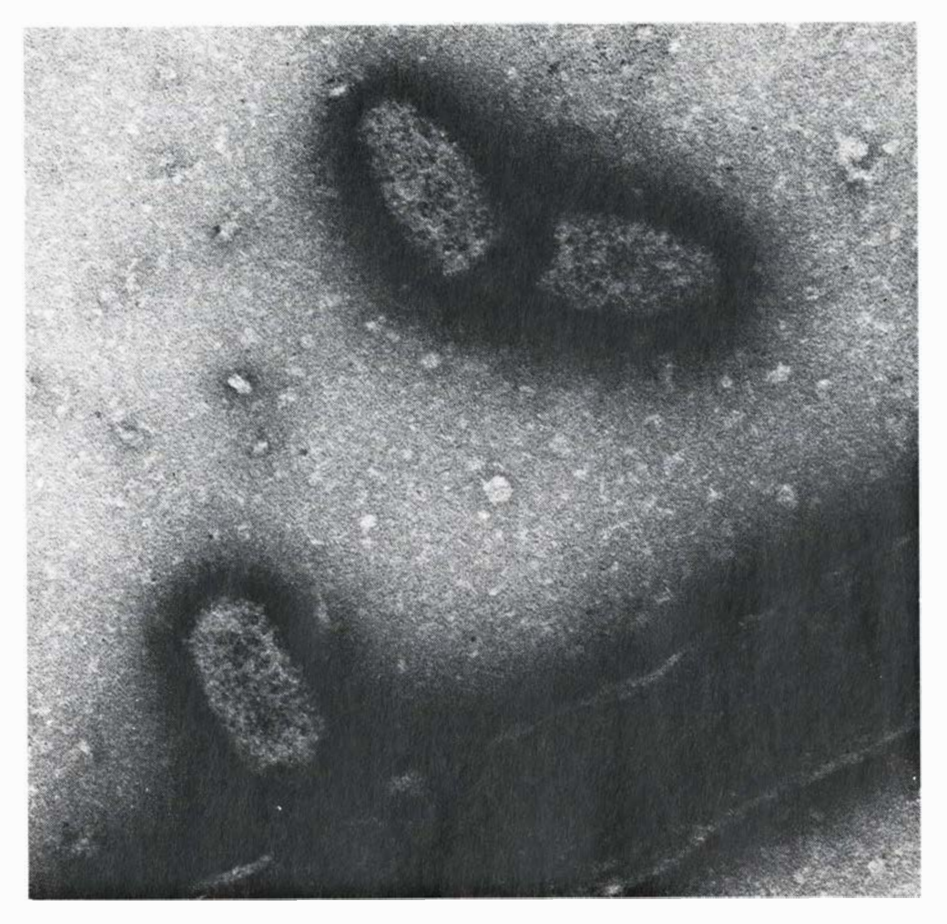
Las partículas víricas penetran por las terminaciones nerviosas y progresan hacia el cordón espinal por el axoplasma (el fluido protoplásmico que circula por los axones de las neuronas, o células nerviosas, del tronco nervioso). El movimiento del virus en los axones es pasivo y avanza a razón de unos tres milímetros por hora. Los axones conducen al cuerpo celular de las neuronas situadas en los ganglios, en las ramas dorsales de los nervios espinales, poco antes de entrar en la médula espinal. A ésta llega la infección, por la citada via; en ella avanza progresivamente a través de las sinapsis neuronales, afectando de paso a los ganglios colaterales que se encuentran a lo largo del cordón espinal. Por éste sigue progresando el virus hasta alcanzar rápidamente el cerebro, donde se multiplica. A partir de aquí, siguiendo los troncos nerviosos, el virus invade todo el cuerpo sin excluir las glándulas salivares, un lugar crucial para la propagación de la enfermedad a otros animales pues en estas glándulas el virus sigue multiplicándose.



posterior con otro detergente (DOC) rompe la partícula, lo que favorece el aislamiento de una nucleocápside compuesta por unas 30 o 35 madejas de ribonucleoproteínas. Mediante un nuevo tratamiento posterior con SDS y 2-ME se digieren los componentes víricos, lo que permite separar la cadena de ácido ribonucleico (ARN) de la proteína de la nucleocápside. El análisis químico nos muestra que la glucoproteína representa el 48 por ciento de la proteína vírica; de éste, el 32 por ciento corresponde a la proteína de la nucleocápside y el 20 por ciento a la envoltura proteica. En la nucleocápside existe otra fracción muy pequeña de proteína no estructural. Lípidos como el colesterol pueden constituir el 1 por ciento del virus.

Para que el ciclo se complete se necesitan semanas o meses. Se desconoce aún cómo o dónde permanece el virus en estado latente e indetectable durante tan prolongado período de incubación. Tampoco se conoce qué factores determinan que la rabia se desarrolle en unos individuos y en otros no, aunque por los síntomas y la causa de la muerte (fallo respiratorio) es evidente que la función neuronal resulta seriamente afectada. Tampoco conocemos qué detiene la enfermedad en la mayoría de los sujetos infectados naturalmente, incluyendo a los que han sido vacunados. En ello parecen estar involucrados anticuerpos neutralizantes, ciertos linfocitos asesinos ("Killer") que atacan a las células infectadas por el virus y el interferón (proteína fabricada por unas células que son estimuladas por varios agentes infecciosos entre los que se encuentra el virus rábico). Queda por dilucidar la interrelación de todos estos factores.

La rabia se extiende por casi todo el



REPRODUCCION DEL VIRUS en una célula infectada, según aparece al microscopio electrónico. El cuerpo oscuro que cruza diagonalmente la parte inferior de la micrografía es la membrana celular. Puede verse cómo una partícula de virus rábico emerge de la membrana; otras dos están por encima.

planeta. Grandes zonas que se hallan a salvo de la misma son: Australia, la Antártida y ciertas islas, entre las que se encuentran Nueva Zelanda, Hawai, Inglaterra e Irlanda y los países escandinavos, excepto las zonas fronterizas de Dinamarca. Durante siglos el principal vector de la enfermedad ha sido el perro, mas por una serie de circunstancias (como cambios ecológicos y la vacunación de los perros), en los últimos años y particularmente desde la Segunda Guerra Mundial éste ha sido reemplazado en muchas partes por otros animales.

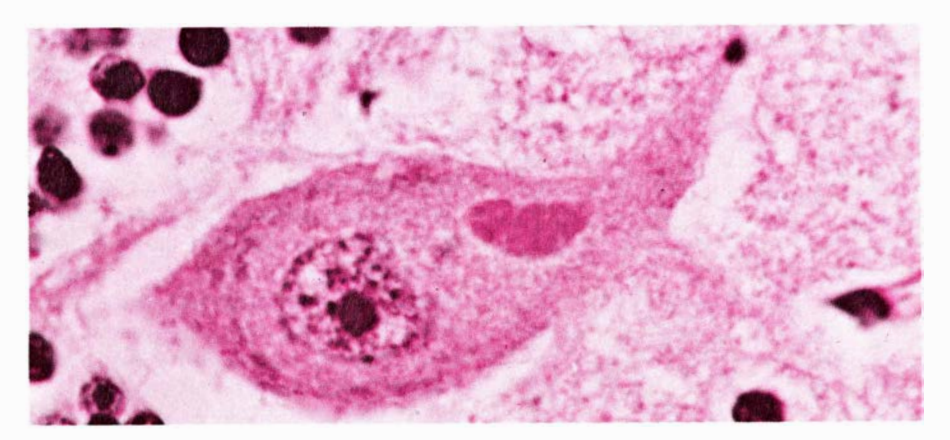
Todos los animales de sangre caliente son susceptibles de sufrir la enfermedad, tanto de forma natural como por inoculación experimental, si bien dicha susceptibilidad varía ampliamente de unas especies a otras. Así, por ejemplo, la mofeta, la zarigüeya y las aves son relativamente resistentes; por contra, zorros gatos y ganado son muy susceptibles. En esa gama, el hombre y el perro ocupan una posición intermedia.

En el perro, el período de incubación suele durar de tres a seis semanas, aunque pueden bastar unos diez días y, en el otro extremo, prolongarse por un año o más. Los primeros signos apreciables son un humor expectante y molesto y un cambio de carácter, seguido de inquietud, dando dentelladas a moscas imaginarias y lamiendo o mordisqueando el lugar de la mordedura. El animal se suele encontrar como asustado y huidizo.

Entre el tercer y quinto día tras la aparición de los síntomas, surge el síndrome paralítico, bien en forma furiosa o muda. En la forma furiosa, el perro está inquieto y tiende a morder los objetos en movimiento; hacia el final la mandíbula inferior suele colgarle y el perro babea. A continuación aparecen la parálisis y las convulsiones. En la forma muda el dueño puede pensar que el perro tiene algún cuerpo extraño alojado en la garganta. En este caso no suele estar irritable y en rara ocasión muerde, se muestra adormilado y con tendencia a esconderse.

En los gatos la enfermedad cursa de forma similar. En el caso del caballo y del ganado los signos consisten en irritación, inquietud y una agresividad anormal. Sin embargo, estos animales no suelen considerarse como vectores de la enfermedad ya que es muy raro que muerdan al hombre o a otros animales.

En los animales salvajes la característica principal de la rabia es la pérdida de temor al hombre y a los grandes animales. Este carácter se hace evidente, particularmente en el caso del zorro, chacal, mofeta y murciélago. Por lo demás, la sintomatología es muy similar a la descrita para el perro. Cuando cualquier animal salvaje ataca y muerde sin provocación, con la posible salvedad de la rata de ciudad, hay motivos suficientes para sospechar de rabia.



LOS CUERPOS DE NEGRI proporcionan un dato fundamental de diagnóstico de rabia. La estructura en forma de pez que se observa en el centro de la micrografía electrónica es una célula del cerebro de un bovino víctima de la rabia. El objeto elíptico y oscuro que se ve en la parte superior derecha de la célula es un cuerpo de inclusión de Negri, denominado así en honor del médico italiano Adelchi Negri, que descubrió en 1903 esta estructura característica. La estructura aparece en el citoplasma de la célula infectada en los últimos estadios de la enfermedad. Estos cuerpos de inclusión pueden no encontrarse o bien confundirse, por lo que actualmente no representan sino un signo más de diagnóstico clínico.

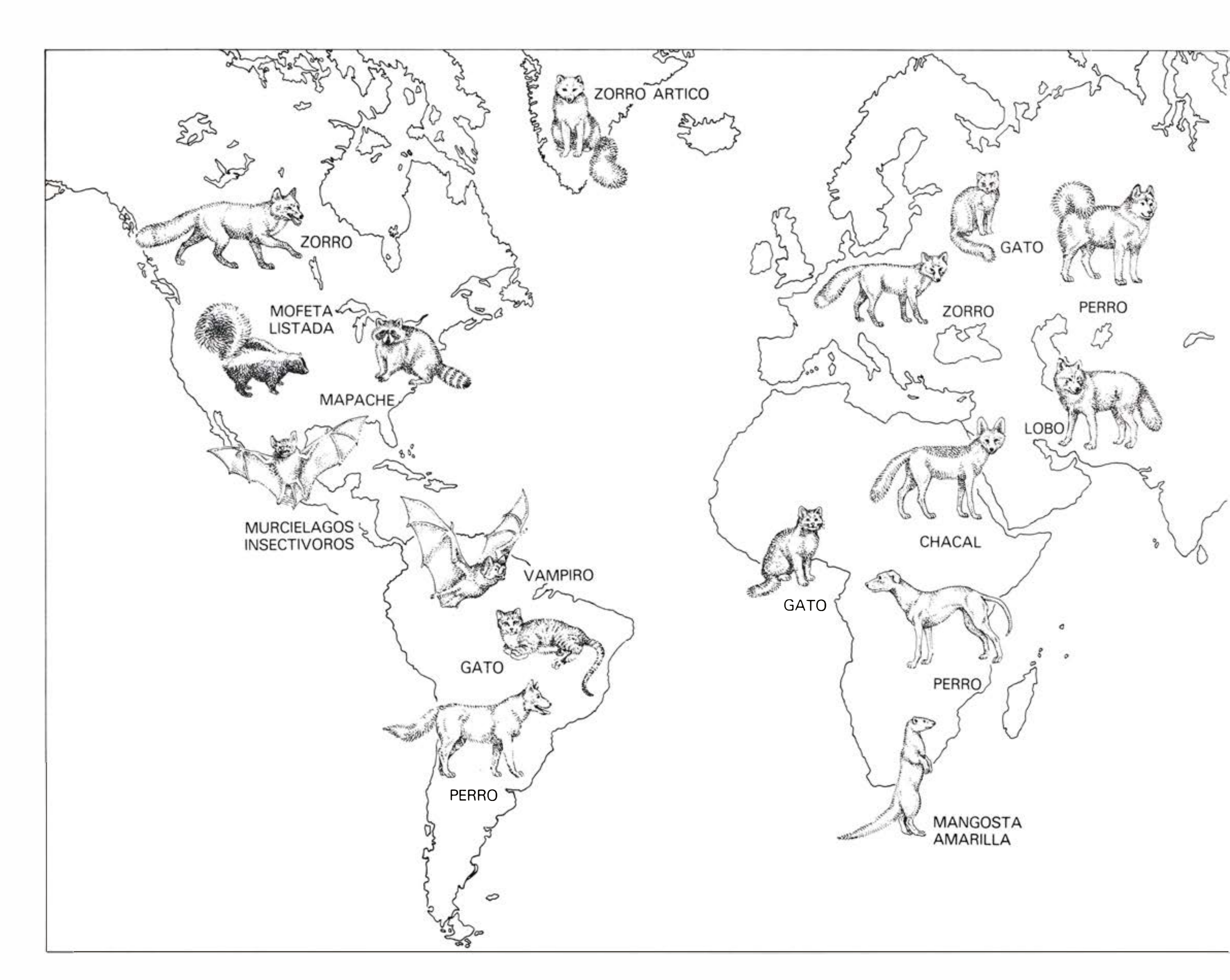
En muchas partes del mundo el mayor problema epidemiológico de la rabia lo constituyen los animales salvajes. A principios del siglo XIX, en la baja California, la rabia en el hombre se asoció a la mofeta manchada (*Spilogale putorius*). Sin embargo, a lo largo de los últimos años otras especies de mofetas

(principalmente las listadas, *Mephitis mephitis*) han estado implicadas en más de la mitad de todos los casos de rabia registrados en los Estados Unidos, habiéndose comprobado que este animal excreta el virus en su saliva incluso 18 días antes de su muerte. Durante cuyo período se torna muy agresiva. Estas características hacen de la mofeta una reserva eficaz y un vector de la rabia tanto para las demás mofetas como para otros animales salvajes.

En Europa, ya a principios del siglo XIX, se encontraron zorros rabiosos, pero no se les dio mayor importancia hasta el siglo XX. Los motivos por los que la incidencia ha aumentado no se conocen. Actualmente, tanto el zorro rojo como el gris constituyen el primer reservorio de la rabia en el continente europeo, así como en toda la región circumpolar norte, que comprende a Groenlandia, parte de Asia y América del Norte. En Europa concretamente, los zorros rabiosos representan un problema de especial gravedad, ya que la enfermedad avanza progresivamente, a razón de unos treinta kilómetros al año, del este hacia el oeste (observándose este fenómeno a partir de la Segunda Guerra Mundial). A través de Polonia y Alemania ha alcanzado ya las costas del Canal de la Mancha, al norte de Francia, habiendo invadido la mayor parte oriental y meridional de dicho país y el resto de la Europa continental.

En las zonas gravemente afectadas se ha llegado a comprobar la muerte de hasta el 50 por ciento de los zorros. En Alemania Occidental se han visto zorros rabiosos cada tres años, de lo que se deduce que una tasa de muerte del 50 por ciento resulta insuficiente para erradicar espontáneamente la enfermedad. Una densidad de población de un zorro por kilómetro cuadrado es suficiente para mantener la rabia. Experimentos realizados en Dinamarca han demostrado que la enfermedad desaparecía cuando la densidad quedaba reducida a un zorro por cuatro o cinco kilómetros cuadrados.

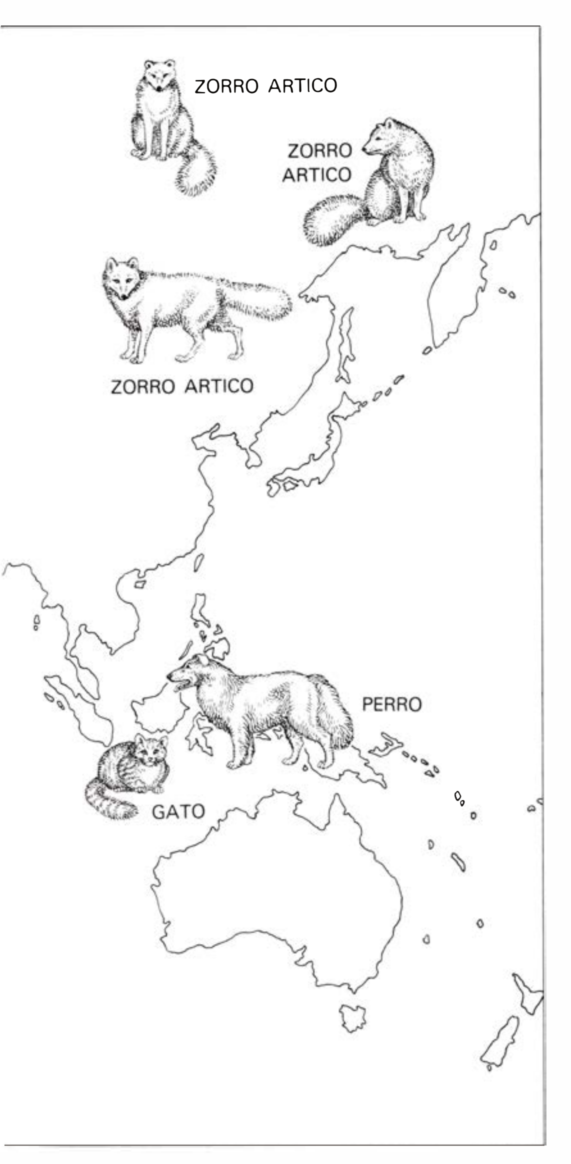
Drácula, la novela ya clásica de Bram Stoker, versa sobre las antiguas leyendas de vampiros y vampirismo. La historia tiene lugar en Transilvania (Rumanía) y se centra en el vampirismo humano, la succión de sangre humana por una persona "inmortal" que abandona su sepulcro durante la noche. Esta superstición se extendió ampliamente por Europa durante el siglo XVIII, en



PRINCIPALES ANIMALES VECTORES de las zonas más importantes en que aparece la enfermedad. En los Estados Unidos las mofetas, sobre todo la listada (*Mephitis mephitis*), representan más de la mitad de todos los casos de rabia diagnosticados en animales; unas 50 especies de murciélagos insectívoros constituyen casi el 15 por ciento. En los lugares en que los perros figuran como la principal fuente de contagio para el hombre, los

particular por Hungría y Servia, hasta el punto de que en Meduenga, un pueblo cercano a Belgrado, en 1732 y después en 1737, una comisión médica compuesta por tres cirujanos militares mandó la exhumación de muchos cadáveres a los que se les aplicó el tratamiento prescrito para los vampiros. Este consistía en clavar una estaca de fresno en el pecho del supuesto vampiro, que solían ser por lo general casos de suicidio. El pasatiempo llegó a generalizarse tanto por toda Europa que en 1823 se decretó una ley en Inglaterra para prohibir específicamente la mutilación del cuerpo de los suicidas por el citado procedimiento.

En Europa se sucedían las leyendas de brujería, murciélagos y vampiros. (Aunque Stoker nunca identificó a Drácula, en sus correrías nocturnas, con un animal, la descripción del espectro sugería la de un murciélago.) Así, a Pedro Mártir de Antequera, el primer obispo del Nuevo Mundo, no le pareció fuera de lugar escribir a principios del siglo XVI: "En muchos lugares, murciélagos no menores que tórtolas vuelan y atacan con furia [a los hombres] y con sus mordeduras venenosas les confieren a los afectados la locura". Observaciones similares sobre mordeduras de murciélagos en el hombre fueron descritas por los primeros colonizadores españoles en la península mexicana del Yucatán y en lo que hoy es Panamá.



gatos están en segundo lugar. Australia, Inglaterra y Escandinavia están libres de rabia. Los animales cumplen largos períodos de cuarentena.

A pesar de la larga tradición popular sobre la mordedura de los murciélagos como causa de rabia, tal relación no se pudo comprobar hasta hace unos 50 años, cuando Joseph L. Pawan observó en Trinidad que los nativos que, a causa del calor, dormían a la intemperie sufrían con frecuencia durante la noche ataques y mordeduras por criaturas oscuras que se movían con rapidez y que Pawan identificó con murciélagos. Las víctimas nunca sentían la mordedura, enfermaban seis semanas más tarde y morían de rabia. Unos 10 años antes ya se había establecido en el sur del Brasil una relación entre la mordedura de murciélagos rabiosos y la rabia paralítica en el ganado.

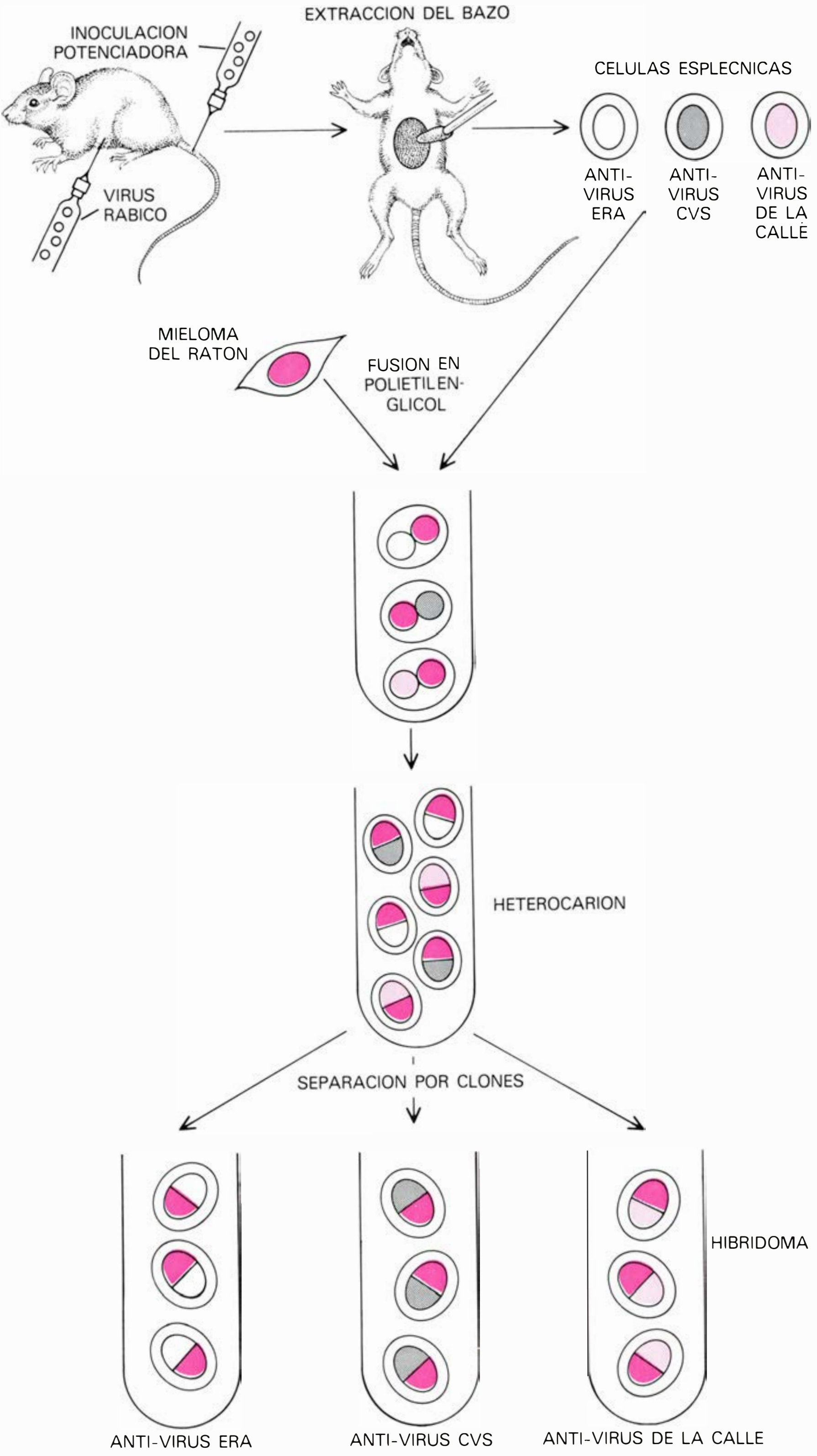
Los vampiros rabiosos (murciélagos) matan al cabo del año cientos de miles de cabezas de ganado en América Latina, donde todos los países, excepto Chile, conocen sus efectos. Aunque la mayoría de los vampiros que han sido inoculados experimentalmente sufren parálisis y mueren al cabo de pocas semanas, algunos permanecen sin síntomas apreciables y esparcen el virus de su saliva durante meses o años. Los murciélagos viven en colonias. Se ha observado que suelen batirse entre ellos, así como con otros que son frugívoros e insectívoros, a los que transmiten la rabia. El primer caso registrado de rabia transmitida por murciélagos al norte de la frontera mexicana se dio en los años 40 y afectó a una jirafa que murió de rabia en el Zoo de San Diego. El episodio se relacionó con una incursión de numerosos murciélagos insectívoros. En los Estados Unidos nunca se consideró en serio la presencia de murciélagos infectados de rabia, hasta que en 1953 se aisló el virus rábico del tejido cerebral de un murciélago insectívoro que, en plena luz del día, había mordido a un niño en Tampa (Florida).

Desde entonces son numerosos los casos descritos de seres humanos mordidos por murciélagos insectívoros. Entre estos se incluyen más de 50 especies en todos los estados de la Unión y en muchos países iberoamericanos. No está claro si los miles de murciélagos insectívoros infectados naturalmente encontrados en Norteamérica durante los últimos 25 años representan un recrudecimiento de la enfermedad o sólo reflejan el descubrimiento de una incidencia de la enfermedad que se ha dado siempre entre los murciélagos.

Hay pruebas de que la rabia de los murciélagos puede transmitirse sin mordeduras. En Texas, un entomólogo estatal que pasó mucho tiempo en la Cueva del Frío, en el condado de Uvalde, contrajo la rabia en 1957 sin haber sufrido mordedura alguna; probablemente la enfermedad se transmitió por inhalación de secreciones de murciélagos infectados. Todos los coyotes, zorros y zarigüeyas enjaulados y protegidos por mallas, expuestos al aire de la cueva, murieron de rabia. Andando el tiempo, se aisló también el virus rábico del aire de la cueva.

Antes del advenimiento de la gnoseología médica moderna el tratamiento de la rabia consistía en una serie de remedios, a veces verdaderamente grotescos. Los antiguos médicos chinos recomendaban, como terapéutica infalible, el cinabrio y el almizcle. Asimismo, en su día se elogió el excremento de aves acuáticas, el ajenjo y las cenizas de hipopótamo. En 1806, el municipio de Nueva York recompensó con 1000 dólares a John M. Crous por el descubrimiento de un "remedio" que consistía en tabletas hechas a base de: quijada de perro pulverizada, lengua desecada de un potro recién nacido y limaduras de cobre de un penique inglés acuñado durante el reinado de Jorge I. En las zonas fronterizas del oeste americano era frecuente que los animales mordidos se trataran con la "piedra de loco", que eran cálculos biliares de un raro ciervo blanco o de una res blanca.

La rabia fue la primera enfermedad humana para la que tuvo éxito una vacuna desarrollada de manera sistemática. Pasteur y sus colaboradores, en 1881, encontraron que el "veneno" de la rabia (el virus) podía obtenerse relativamente puro a partir del cerebro y médula espinal de un animal infectado. Un avance crucial se registró con el descubrimiento de que por pases sucesivos en conejo del virus de la calle, que tenía un período de incubación largo y variable, se obtenía una forma fija de virus con un período de incubación corto y constante de cinco a siete días. Este tipo de



LA TECNICA DE LOS ANTICUERPOS MONOCLONALES proporciona un método altamente específico para diferenciar entre sí las cepas de virus rábico. Los ratones se inmunizan con virus rábico o con componentes aislados de los mismos, tales como una glucoproteína o una nucleocápside. El antígeno se inocula intraperitonealmente y, pasadas tres o cuatro semanas, se hace otra inoculación intravenosa para potenciar la respuesta. Tres o cuatro días más tarde, se extrae el bazo para obtener una suspensión de linfocitos; ésta se mezcla, en presencia de polietilenaglicol, con células cultivadas de mieloma de ratón, deficientes en el enzima hipoxantín-guanín-riboxiltransferasa. Las células fundidas, llamadas heterocariones, se colocan en un medio de cultivo que impide el crecimiento de las células mielomatosas. Dado que los linfocitos no pueden crecer indefinidamente en cultivo, las colonias que aparecen en las placas de cultivo están constituidas sólo por células híbridas con rasgos de linfocitos y de células mielomatosas. Estos cultivos, llamados hibridomas, se utilizan mediante tratamiento adecuado, para la producción de cultivos hijos procedentes de una sola célula hibridoma (cultivos monoclonales). Seguidamente se examinan los clones para detectar la presencia de anticuerpos específicos frente a cepas de virus rábico como las "ERA", "CVS" y virus de la "calle". Un hibridoma que produzca un determinado anticuerpo antirrabia se puede cultivar y mantener por tiempo indefinido. (P. Wynne.)

virus resultaba eminentemente apropiado para la preparación de la vacuna.

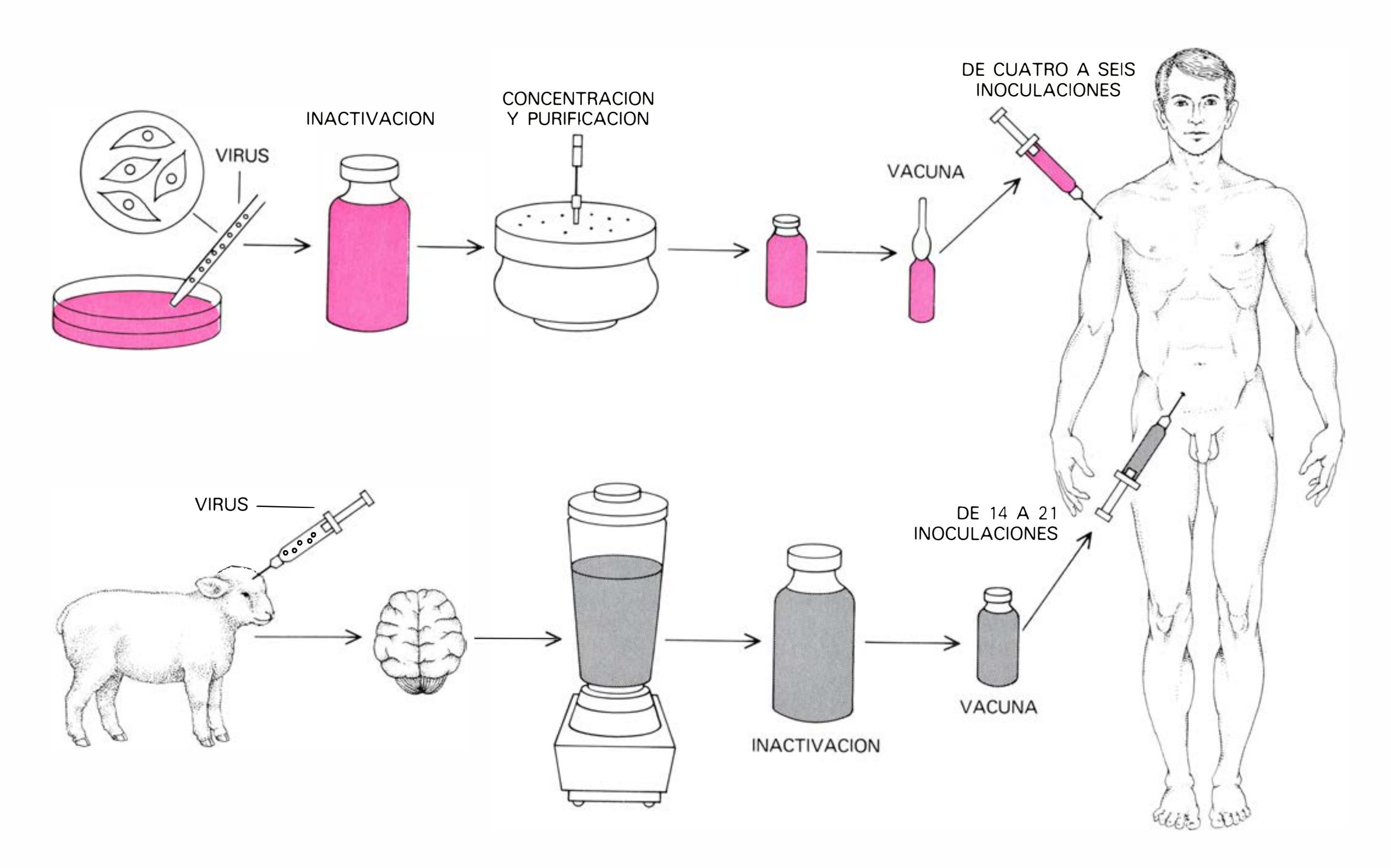
El éxito que tuvo Pasteur con su vacuna en perros le impulsó en 1885 a probarla con un tal Joseph Meister, quien posteriormente sería conserje del Instituto Pasteur. A pesar de que Meister había sufrido graves mordeduras de un perro rabioso, sobrevivió. Esta prueba facilitó el camino para aumentar los ensayos en el hombre. En general los casos se resolvieron satisfactoriamente.

En este sentido, los logros de Pasteur se debieron en parte a un experimento realizado por su colega Émile Roux quien había puesto en una estufa de incubación un frasco que contenía la médula espinal de un conejo muerto de rabia, al que se le había provocado experimentalmente la enfermedad. Roux estaba haciendo ensayos para averiguar cuánto tiempo lograba sobrevivir el virus a la temperatura de 37 grados Celsius (temperatura del cuerpo). Pasteur modificó la técnica adicionando potasa como elemento deshidratante y manteniendo el frasco a la temperatura del laboratorio.

Su idea genial fue preparar una serie de 21 a 28 inoculaciones a partir de una suspensión de médula espinal, comenzando por la que había tenido en desecación durante 14 días, cuando el virus resultaba ya inocuo, y continuando con las suspensiones de médula cuyos tiempos de desecación habían ido disminuyendo desde los 13 hasta los 3 días. Las últimas suspensiones desencadenaban la rabia en los conejos, pero las primeras habían conferido suficiente inmunidad como para proporcionar protección contra la enfermedad.

Desde entonces se han introducido una serie de modificaciones en el método de Pasteur, la mayoría de ellas encaminadas a mejorar la calidad de la vacuna. Sin embargo, el principio de la vacuna, basado en la suspensión del tejido nervioso, se sigue manteniendo aún en la mayoría de países en vías de desarrollo, en razón de su fácil producción y bajo coste.

Con todo, el procedimiento de Pasteur no estaba exento de complicaciones. La peor de ellas era que el virus fijo conservaba la capacidad de producir la rabia y de hecho así ocurría cuando la vacuna no se preparaba adecuadamente. El mismo Pasteur fue acusado en 1886 de ocasionar la muerte con su vacuna a un niño de 10 años. El denunciante de Pasteur era Georges Clemenceau, que posteriormente sería por dos veces primer ministro de Francia. En la década de 1860 Clemenceau practicó la medi-



PROCEDIMIENTOS VACUNALES comenzando *(parte inferior)* por la técnica clásica iniciada por Louis Pasteur. Se inocula el cerebro de un animal como la oveja con una cepa de virus fijo de la rabia. Cuando el animal presenta síntomas de enfermedad, se sacrifica y se le extrae el cerebro, que se tritura en un homogenizador. Seguidamente se inactiva el virus por procedimientos físicos o químicos. El tratamiento de una persona expuesta a la rabia consiste en una serie de 14-21 inoculaciones de virus inactivado bajo la piel del abdomen. Este tratamiento es doloroso y a veces va acompañado de reacciones indeseables. La nueva vacuna (*serie superior de la figura*) se fabrica a partir del cultivo de tejido infectado con una cepa de virus fijo adaptado de rabia. El virus se inactiva y el producto resultante se concentra y purifica. Esta vacuna es más potente que la de Pasteur.

cina durante algún tiempo en Nueva York y enseñó francés en un colegio femenino de Stamford, Connecticut. Fue partidario incondicional de la teoría de la generación espontánea (a la que Pasteur se oponía) y perseveró en su idea hasta su muerte de 1929. En el juicio, Pasteur recibió el apoyo de Paul Brouardel, director del laboratorio, y de Roux (a pesar de su resentimiento hacia Pasteur por haber copiado éste su procedimiento de desecado). Brouardel y Roux comprendieron que un fallo contra Pasteur supondría un retroceso de décadas para la ciencia médica francesa y expondría al laboratorio a posibles litigios. Pasteur fue absuelto y su procedimiento vacunal pudo utilizarse ampliamente a través de los institutos Pasteur establecidos en otros países.

A pesar de todo, algunos pacientes murieron de rabia aun después de haber recibido el tratamiento completo de Pasteur. Por ello, en 1889 Victor Babes y M. Lepp desarrollaron una nueva vía en el tratamiento humano de quienes habían quedado expuestos a la rabia. Inmunizaron animales con el virus rábico y obtuvieron de su sangre el suero que contenía anticuerpos frente al virus. Este suero era administrado entonces al mismo tiempo que la vacuna de Pasteur. El procedimiento no se generalizó en aquel tiempo debido a que los resultados obtenidos fueron un tanto equívocos.

En la década de los 40 Karl Habel, de los National Institutes of Health, reavivó el interés por la técnica del suero, y, gracias a estudios posteriores llevados a cabo por uno de nosotros (Koprowski), en 1950 fue recomendado por el Comité de la Rabia de la Organización Mundial de la Salud (OMS) un ensayo en humanos de un suero antirrábico obtenido de caballos. La oportunidad surgió en 1954 cuando un lobo rabioso mordió a 29 personas y varios animales domésticos en un pueblo del Irán. Se estableció un plan experimental en el que con un grupo de víctimas se siguió el esquema normal de vacunación y al otro se le administró el suero, además de la vacuna. De trece que recibieron el tratamiento doble sólo murió uno, mientras que entre los que recibieron el tratamiento estándar, de cinco con heridas similares a las del grupo anterior murieron tres y las once restantes, cuyas heridas eran más leves, sobrevivieron.

Hoy día está generalizada la aplicación de gammaglobulinas (concentrado de la fracción proteica del suero que contiene los anticuerpos) junto con la vacuna. Sin embargo, dado que muchas vacunas corrientes tienen una potencia escasa, la inmunidad a corto plazo proporcionada por el suero puede interferir con el efecto inductor de la vacuna en la producción de anticuerpos, lo que obliga a buscar vacunas más potentes para superar dicha interferencia. Otros motivos que inducen a intentar el desarrollo de vacunas más potentes son: acortar el tiempo de dos o tres semanas de inoculaciones diarias requerido por las vacunas corrientes y eliminar sus efectos colaterales, que a veces llegan a ser tan graves como para causar parálisis e incluso la muerte.

En la década de los 60 Wiktor, Stanley A. Plotkin y nosotros emprendimos intensivos estudios de laboratorio al objeto de desarrollar una vacuna preparada con virus crecidos en células embrionarias humanas normales. Tras casi 10 años de investigación se consiguió una vacuna de gran potencia que protegía al mono inoculado con virus rábico virulento, incluso con una sola inocula-

ción vacunal administrada a las pocas horas de haberle atacado el virus virulento.

La nueva vacuna se sometió a un riguroso control en las pruebas realizadas en humanos en 1975 y 1976 en el Instituto Pasteur de Teherán, bajo la dirección de M. Bahmanyar, quien había participado en las primeras pruebas con personas mordidas por lobos. Los últimos ensayos comprendían dos episodios: uno de mordeduras de lobo que afectaba a nueve personas y otro que incluía seis casos de mordedura de perro, en el que estaban implicadas treinta y seis personas. Se demostró que, en todos los casos, los animales mordedores estaban rabiosos. Los individuos expuestos recibieron suero y un total de sólo seis dosis de la nueva vacuna, en lugar de las 21 inoculaciones requeridas por el esquema vacunal clásico. El tratamiento de algunas víctimas no comenzó hasta pasados tres días del accidente (en un caso a los 14 días). Por experiencia se sabía que una dilación de ese orden podría inducir la muerte rábica en algunos casos; sin embargo, ninguna de las 45 personas llegó a morirse. Esta vacuna se emplea, en la actualidad, de forma rutinaria en muchos paises.

Una de las medidas preventivas más importantes en la rabia, tras la exposición, consiste en tratar lo más pronto posible la herida con soluciones destructoras de virus como agua con jabón, alcohol, solución yodada y desinfectantes a base de amonio cuaternario. Si se dispone de suero antirrábico es muy recomendable hacer una infiltración con el mismo alrededor de la herida. Si las circunstancias lo permiten debe retrasarse al máximo la sutura de la herida.

Aunque se han registrado casos de recuperación una vez manifestada la enfermedad, esto resulta extremadamente difícil. En las dos últimas décadas sólo se han registrado la supervivencia de dos pacientes, que se trataron controlando sus funciones vitales (el sistema nervioso y cardio-pulmonar) e interviniendo rápidamente para restablecerlas, una vez que la enfermedad alcanzó el estadio crítico de excitación y coma. Salvo estas raras excepciones, siguen siendo válidas las observaciones que en su día hizo Moisés Maimónides (1198) en su "Tratado sobre Venenos": "Todo lo mencionado en la literatura contra la mordedura de un perro rabioso puede ser útil, siempre y cuando se aplique antes de que la enfermedad se manifieste. Si esto último ocurre, aún no sé de nadie que haya salido con vida".

Otra vía de lucha contra la rabia consiste en intentar su control en los animales. Muchas de las islas libres de rabia lo deben, en gran parte, a observar una estricta cuarentena de hasta varios meses, en todos los perros, gatos y otros carnívoros relacionados que entran en ellas. Asimismo, existen vacunas eficaces y baratas que protegen los animales frente a la rabia. Una inoculación preventiva es suficiente para proteger a la mayoría de los animales por un período de hasta tres años. El control de la rabia en los perros y gatos de las ciudades se consigue fácilmente mediante programas adecuados de vacunaciones preventivas y evitando la presencia de animales callejeros sin dueño. La vacunación preventiva también se puede utilizar para proteger a la ganadería.

Por el contrario, el control de la rabia en los animales salvajes resulta extremadamente difícil. Los esfuerzos realizados en algunos países para menguar los casos de zorros rabiosos, haciendo disminuir la densidad de población de los mismos, sólo han demostrado una eficacia temporal y efímera. En Europa se ha ensayado la vacunación oral de zorros utilizando cebos adecuados (cabezas de pollo conteniendo cápsulas con vacuna de virus atenuado). Es un procedimiento cuestionable porque aún se sabe poco sobre el comportamiento de tales cepas de virus bajo las condiciones ambientales y tras pases reiterados de zorro a zorro. El virus podría recuperar su virulencia.

Mayor dificultad presenta el control de los murciélagos rabiosos. Las medidas tomadas en Iberoamérica para reducir la población de los vampiros en los países más afectados de rabia están muy lejos del éxito, ya que los procedimientos utilizados (cebos con anticoagulantes sanguíneos) sólo son factibles a escala local e inaplicables con otras especies de murciélagos. La medida más común y adecuada de protección frente a la rabia vehiculada por los murciélagos se basa en la vacunación preventiva del ganado.

Aunque la rabia es una de las enfermedades más antiguas reconocidas en el hombre, quedan por esclarecer muchos de sus factores (y sus interrelaciones) como: la latencia y el largo período de incubación de los virus, el papel jugado por los anticuerpos, por la inmunidad linfocitaria y por el interferón. Dilucidar tales cuestiones no sólo será importante para conocer más y luchar mejor contra la enfermedad de la rabia, sino también para otras muchas infecciones del sistema nervioso central que hoy día permanecen como un misterio tanto desde el punto de vista médico como biológico.

Enzimas alostéricos

La organización oligomérica de los enzimas alostéricos sirve de base tanto al fenómeno de cooperatividad en la unión y en la catálisis como a su modulación por efectores. El mismo diseño parece repetirse en otras proteínas reguladoras

Eduardo Cadenas

La biología progresó muy lentamente hasta que la química logró cierto grado de desarrollo, porque los organismos, que constituyen su objeto de estudio, son realmente máquinas químicas. Más acertado sería decir que, debido a esta circunstancia, existió un notable retraso en el avance de ambas ciencias hasta el momento en que se estimularon mutuamente. Y es que el hombre, que entiende con facilidad los mecanismos o máquinas mecánicas, está mal dotado para comprender las máquinas químicas, o quimismos, por su propia fisiología, aunque él mismo sea una de estas últimas. Además es la propia índole de uno y otro tipo de dispositivo lo que determina nuestra actitud, con independencia de nuestras limitaciones. Unos y otros requieren energía para su funcionamiento. Es precisamente esto lo que les confiere el carácter de máquinas. Igualmente cierto es que ambos están sometidos a regulación. Termodinámica y cibernética se aplican por igual a su descripción, pero compete a la biología molecular señalar las notables diferencias que presentan tanto en montaje como en operación. Las partes que constituyen un mecanismo tienen un tamaño y gozan de una permanencia que son del mismo orden de magnitud que los que caracterizan al conjunto. En un quimismo, en cambio, las piezas son moléculas y participan en reacciones químicas en que se transforman, siendo reemplazadas por otras. La operación de una máquina mecánica es esencialmente espacial y sólo exige la comprensión de diversos movimientos. En claro contraste, la función de una máquina química es netamente temporal y requiere la definición de variadas acciones.

Haciendo honor a su denominación, cuando inicia su andadura la química fisiológica se esfuerza por encontrar la explicación de la función biológica a través del conocimiento de la estructura de las moléculas responsables. Durante décadas, el estudio de la naturaleza de los metabolitos sirvió para proponer los mecanismos por medio de los cuales tienen lugar sus transformaciones en la célula. Pero este enfoque no fue el único empleado. Con las primeras investigaciones bioquímicas comenzó ya a tomar cuerpo una estrategia mucho más elaborada que ésta netamente química. Conforme a ella, se parte de una descripción de las funciones vitales en términos de acciones. Después se hace responsable de cada una de las acciones a una entidad concreta cuya naturaleza se desconoce. Para explicar su comportamiento se propone un modelo al que se añaden nuevas facetas a medida que se descubren nuevos detalles sobre la acción. Finalmente, se aborda por métodos directos el estudio de la entidad responsable con el fin de contrastar su comportamiento con el que el modelo le había asignado, en base a inferencias de carácter indirecto. Existen ejemplos en que los avances trascendentales en el conocimiento de la acción fueron guiados por la enunciación del modelo. En otras ocasiones en cambio, el modelo no se propone hasta que se dispone de un rico acervo de datos sobre la acción.

El modelo de la acción enzimática, propuesto en 1913 por Michaelis y Menten, consiguió alcanzar un increíble grado de detalle antes de que, en la década de los años 60, el estudio de la estructura en el espacio de la molécula enzimática permitiera contrastarlo. Sólo ahora comenzamos a disponer de algunos receptores aislados y, sin embargo, por los mismos años en que se propuso el modelo enzimático, ya iba tomando forma el modelo del receptor (Ehrlich). Primero sirvió para encauzar los estudios sobre sustancias activas exógenas y más tarde pudo acoger otras que, como las hormonas y los neurotransmisores, tienen caracter endógeno. En claro contraste se encuentran modelos como el del operón y el alostérico que surgieron como resultado de la acumulación de observaciones sobre determinadas acciones. Durante décadas se reunieron datos referentes a la inducción y represión de la síntesis de determinados enzimas, tanto en sus aspectos fisiológicos como genéticos, antes de que Jacob y Monod formulasen el modelo del operón en 1961. De la misma forma, cuando Monod, Wyman y Changeux formulan su modelo alostérico en 1965, ya hacía más de medio siglo que se venía estudiando la acción de la hemoglobina como transportadora de oxígeno en la sangre.

Sea como fuere, se puede decir con justicia que los diversos modelos que hoy guían la investigación bioquímica constituyen verdaderos monumentos del ingenio humano. Todos ellos poseen una gran originalidad en su planteamiento y una enorme fecundidad en su desarrollo, que en muchos casos desborda los fenómenos que se propusieron explicar. En este aspecto destaca el modelo alostérico, que a través de un quimismo bien pudiera servir de fundamento para los más diversos sistemas de control biológico.

El estímulo para los estudios sobre regulación biológica surgió con motivo de observaciones sobre la actividad de sustancias exógenas o endógenas sobre el organismo (para cuya explicación se propuso el modelo del receptor). Pero los avances más espectaculares se iniciaron al acometer la investigación de las vías anabólicas en los microorganismos. Muy pronto se descubrió que el flujo a través de las rutas de biosíntesis está controlado por determinados metabolitos, cuya acción es análoga a la de las sustancias activas. Los receptores para estos metabolitos son precisamente ciertos enzimas cuya actividad se modifica, como consecuencia de la unión de los enzimas con los metabolitos, de una forma altamente significativa. En 1954, Novick y Szilard reseñan que la síntesis de un precursor del triptófano es inhibida en *E. coli* por este aminoácido. Señalan entonces el alto significado regulador que este hecho entraña. Trabajando con un mutante del mismo microorga-

nismo incapaz de sintetizar treonina, Umbarger logra establecer en 1956 que el aminoácido isoleucina corta su propia producción a partir de treonina, al inhibir el primer enzima de la vía correspondiente. Se aprecia con nitidez que la inhibición por el producto final es un caso típico de retroacción negativa [*véase la ilustración inferior*]. Su utilidad está clara. Impide el derroche que para la célula supondría continuar la síntesis de un producto cuando ya se halla a disposición de las vías que lo utilizan. Su eficacia es igualmente evidente. La acción se centra sobre el primer enzima de la vía. De esta manera se evita no sólo la formación del producto final, sino también la acumulación de los productos intermediarios. Aunque la inhibición por el producto final se descubrió en los procariotas, no es privativa de este tipo de organización celular. Pocos años después de su detección en bacterias se encontró en organismos encariotas. Tampoco es exclusiva de las vías anabólicas. El mismo tipo de fenómeno se observa en toda clase de vías metabólicas.

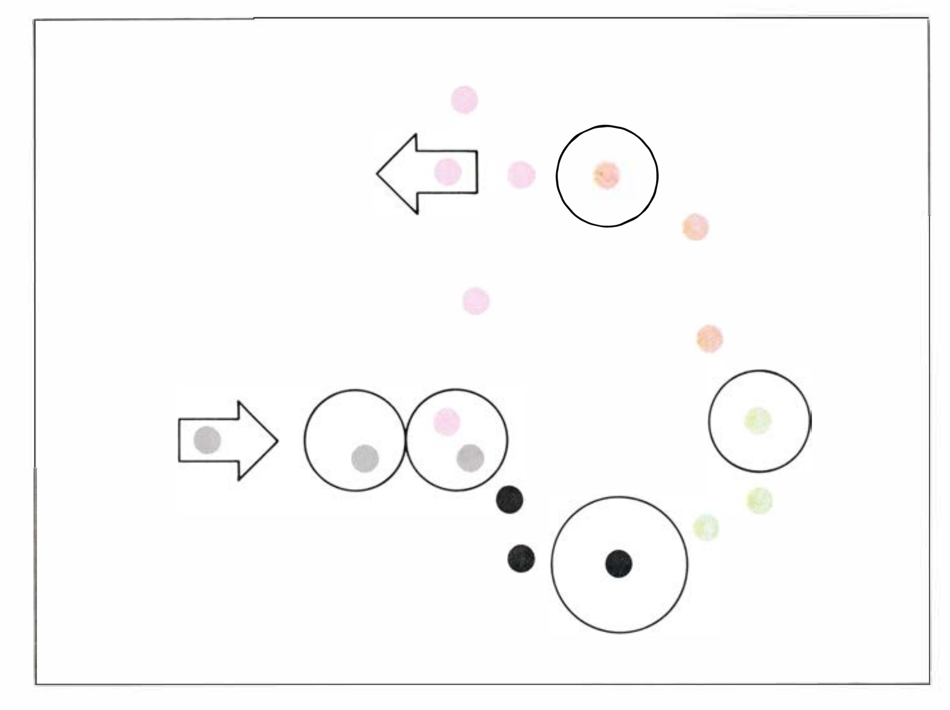
Aunque entre unos sistemas de control y otros existen diferencias, en todos se repite la misma pauta. En cada tramo de una vía metabólica existe un enzima regulador sensible a determinado metabolito, que representa el papel de señal química, que informa en cada momento del estado en que se encuentra el balance entre la productividad de la vía y las necesidades celulares. La unión del metabolito o efector con el enzima regulador provoca una respuesta que consiste en una modificación del flujo por la vía y una variación en los niveles de sus compuestos intermediarios. La respuesta se genera mediante modificación de los parámetros cinéticos del enzima como consecuencia de la unión con el efector y resulta, según los casos, en un aumento o en una disminución de la actividad del enzima o en una alteración de su especificidad, pero en definitiva conduce a un cambio en la productividad de la vía. Una de las características más notables de los enzimas reguladores es que responden a sustancias cuyas moléculas no muestran ninguna semejanza con las de sus sustratos. En base a esta característica, Monod, Changeux y Jacob propusieron para estos enzimas el término de alostéricos (1963) que pretende indicar que deben poseer centros específicos para sus efectores, distintos de los que tienen para sus sustratos (centros activos).

Al investigar la cinética de los enzimas alostéricos se descubrió una notable faceta de su comportamiento. La representación de la velocidad de la reacción frente a la concentración de sustrato, que normalmente proporciona una hipérbola, exhibía una curiosa curva sigmoide. De hecho, un enzima regulador muestra generalmente cinética sigmoide sólo para uno de sus sustratos, mientras que es hiperbólica para el otro. Es como si el enzima fuese capaz de operar conforme a dos patrones diferentes. Esta dualidad de comportamiento también se manifiesta cuando la presencia de un efector convierte, en hiperbólica, la cinética sigmoide del enzima.

Muchos años antes de que se descubriera cómo actúan los enzimas alostéricos ya se conocían las características de la unión del oxígeno con la hemoglobina (véase "La estructura de la hemoglobina y el transporte respiratorio", M. F. Perutz, INVESTIGACION Y CIENCIA, febrero, 1979). Por eso no es de extrañar que se señalara la coincidencia entre la forma de la cinética de un enzima alostérico y la curva sigmoide de saturación con oxígeno de la proteína transportadora. Esta semejanza pudo servir para interpretar el significado de este tipo de cinética en la regulación metabólica y también para hallar su fundamento a nivel molecular.

El carácter sigmoide de la curva de saturación de la hemoglobina explica cómo esta proteína puede descargar la mayor parte del oxígeno que transporta al llegar a los tejidos, donde la presión parcial del gas no es mucho menor que la que prevalece en los pulmones. En la curva de saturación existe un tramo, centrado en los valores de presión usuales para el organismo, en que variaciones relativamente pequeñas de presión determinan una notable liberación de oxígeno. Análogamente, la cinética del enzima alostérico garantiza una respuesta máxima a la variación en la concentración de sustrato dentro de una determinada gama de valores. En particular, hace posible que la actividad enzimática se reduzca notablemente sin necesidad de que la concentración de sustrato tenga que alcanzar para ello valores muy bajos. A través de la acción de un efector, se puede trasladar el tramo de máxima sensibilidad hacia valores más bajos o más altos de la presión parcial de oxígeno, en el caso de la hemoglobina o de la concentración de sustrato, en el de un enzima alostérico. Es en el proceso de desplazamiento hacia concentraciones menores cuando la curva deja de ser sigmoide y pasa a ser hiperbólica.

La coincidencia de varios grupos hemo (cuatro concretamente) en la molécula de hemoglobina capaces de unirse al oxígeno ha estado presente en todos los intentos para justificar la forma sigmoide de la curva de saturación. Primero fue la ecuación de Hill que supuso la unión simultánea con varias moléculas de oxígeno de la molécula de hemoglobina. Después se introdujo la ecuación de Adair que, en contraste con la de Hill y para corregir sus limitaciones, se basaba en la unión de la molécula de hemoglobina con cuatro moléculas



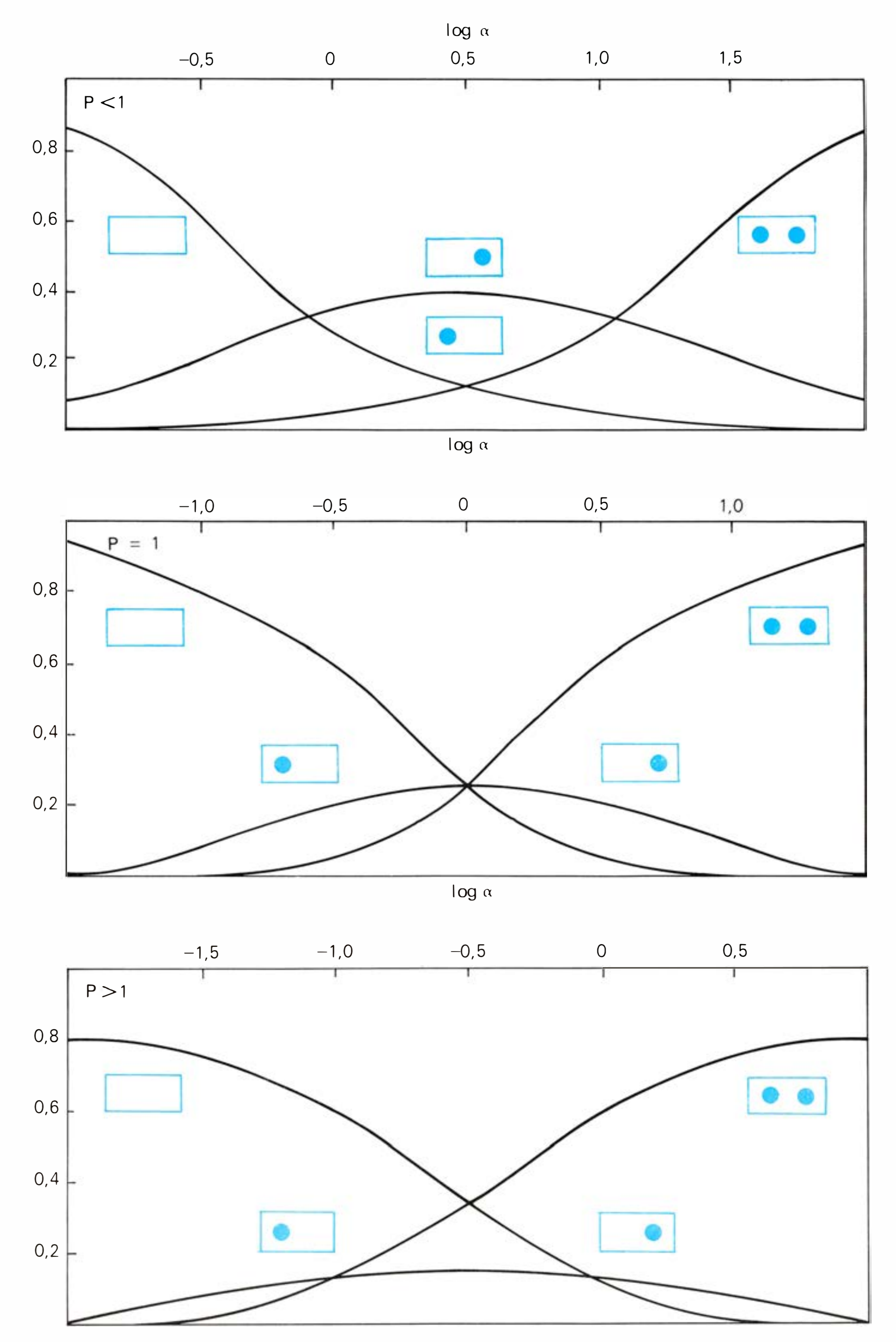
INHIBICION POR EL PRODUCTO FINAL del primer enzima de una vía biosintética. En la vía intervienen varios enzimas que transforman el sustrato inicial (*gris*) en producto final (*rojo*). Aunque la acción de cada enzima supone sólo un retoque de la estructura del sustrato, su actuación en serie determina que entre el producto final y el sustrato inicial no exista apenas semejanza. Por ello, el enzima regulador, que aquí es un dímero, debe poseer centros de unión específicos para el efector (*rojo*).

de oxígeno de forma escalonada. Con esta ecuación como marco, Roughton y colaboradores consiguieron establecer que la unión con cada molécula de oxígeno hace que sea más fácil la unión de la hemoglobina con la siguiente. Otro tanto debe ocurrir en la unión de un enzima alostérico con su sustrato para el que la molécula proteínica debe poseer también varios centros de unión equivalentes.

La unión de una sustancia de peso molecular bajo, ligando, con una proteína es un fenómeno de gran significado biológico. La interacción se localiza en un área limitada de la molécula proteínica que se denomina centro de unión y está caracterizada por una notable especificidad. Se produce a través de una reacción química reversible, cuya constante de equilibrio permite estimar la afinidad del ligando por su centro. Como consecuencia de la reacción, el centro pasa a estar ocupado o saturado por el ligando y la proteína exhibe una variación en sus propiedades. Puesto que la proporción de centros ocupados varía con la concentración del ligando, un número de propiedades de la proteína, entre las que se cuentan algunas de importancia fisiológica, resultan ser función de la concentración del ligando. La mayoría de las proteínas posee varios centros de unión equivalentes por molécula para la unión con un ligando específico. Este hecho en muchos casos sólo implica que la concentración de centros es superior en un pequeño factor a la concentración de proteína. Para algunos, en cambio, el ensamblaje de varios centros de unión en una misma molécula tiene consecuencias realmente sorprendentes.

En principio, por ser equivalentes, todos los centros de unión en una molécula de proteína tienen la misma probabilidad de ocupación. Sin embargo, la situación cambia tan pronto como se van saturando algunos. A partir de entonces, la probabilidad de que un centro se ocupe no es la misma si se halla en una molécula con todos los centros libres, que si está en una molécula con algún centro que ya esté saturado con ligando. De este modo surgen diferencias en la probabilidad de ocupación de los centros entre las moléculas de la proteína. En unos casos se favorece la saturación cuando ya hay centros ocupados en la molécula, en otros ocurre todo lo contrario. Todo depende de qué proteína se trate y de la naturaleza del ligando, siendo normal que una misma proteína presente un comportamiento con un ligando y el contrario con otro. Cuando la saturación de un centro en una molécula de proteína favorece o dificulta la unión con otro, se dice que hay cooperatividad en la unión y que es, respectivamente, positiva o negativa.

Todas las moléculas de una proteína con dos centros de unión equivalentes por molécula para un ligando son iguales en su ausencia: todas tienen sus centros libres. En presencia del ligando en cambio, existen cuatro clases de moléculas proteínicas. Las que tienen sus dos centros libres y las que muestran sus dos centros ocupados y además las que tienen uno u otro de los centros saturado, mientras el otro está libre. La proporción de las moléculas de la proteína que se encuentra en cada una de las cuatro formas depende de la concentración



LA CONCENTRACION DE SUSTRATO influye sobre el reparto de las moléculas de un enzima, con dos centros de unión por molécula, entre sus cuatro formas posibles. La concentración de cada una de las formas se expresa referida a la de la forma con sus centros libres que se toma como unidad: α es la concentración del sustrato tomando como unidad "natural" su afinidad por el enzima, cuando el sustrato se encuentra a baja concentración: ρ representa la relación entre las afinidades del enzima por el sustrato cuando éste se encuentra a alta y baja concentración, porque se supone que, en principio, la afinidad no es la misma a esos dos niveles de concentración. A partir de las concentraciones relativas es fácil calcular la proporción de las moléculas en cada una de las formas. Las correspondientes expresiones permiten representar la proporción de cada forma en función de α o mejor de log α (para de esta forma condensar en poco espacio un amplio margen de concentraciones). La forma en que se cortan las líneas que se obtienen permite detectar si existe cooperatividad o no, e incluso establecer si la cooperatividad en la unión es positiva o negativa. Las gráficas corresponden a ρ = 0,1; ρ = 1 y ρ = 10.

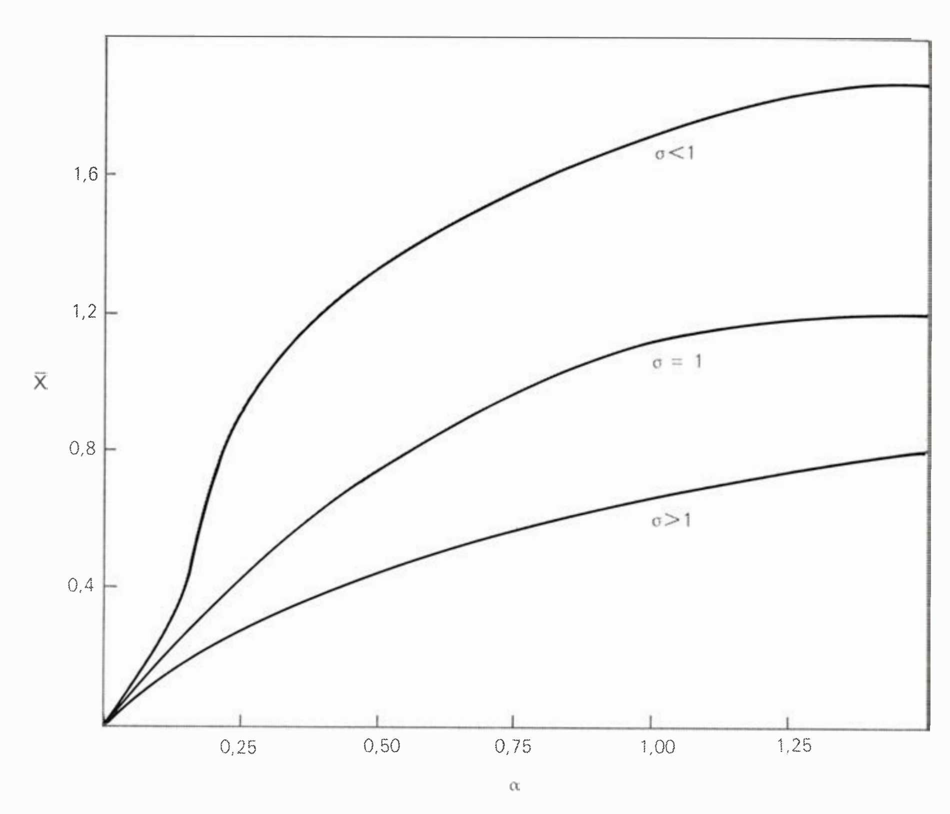
del ligando y varía de modo característico conforme a una función de la concentración del ligando en que aparecen, como parámetros, su afinidad a baja concentración y la relación de afinidades a alta y baja concentración.

Si se representa la proporción de cada una de las formas de la proteína frente a la concentración de sustrato se obtendrá un conjunto de curvas diferente para cada pareja de valores que tengan los parámetros correspondientes. En cambio, si la concentración del ligando se expresa en unidades "naturales" refiriéndola a su afinidad a baja concentración, las curvas obtenidas serán las mismas para todas las proteínas cuya relación de afinidades a alta y baja concentración del ligando coincidan, sin que importe cuál sea la afinidad a baja concentración. Por otro lado, la disposición de los puntos de corte entre las curvas obtenidas en la representación permitirán diagnosticar si hay cooperatividad en la unión y, en caso de haberla, determinar cuál es su signo [*véase la ilustración de la página 108*].

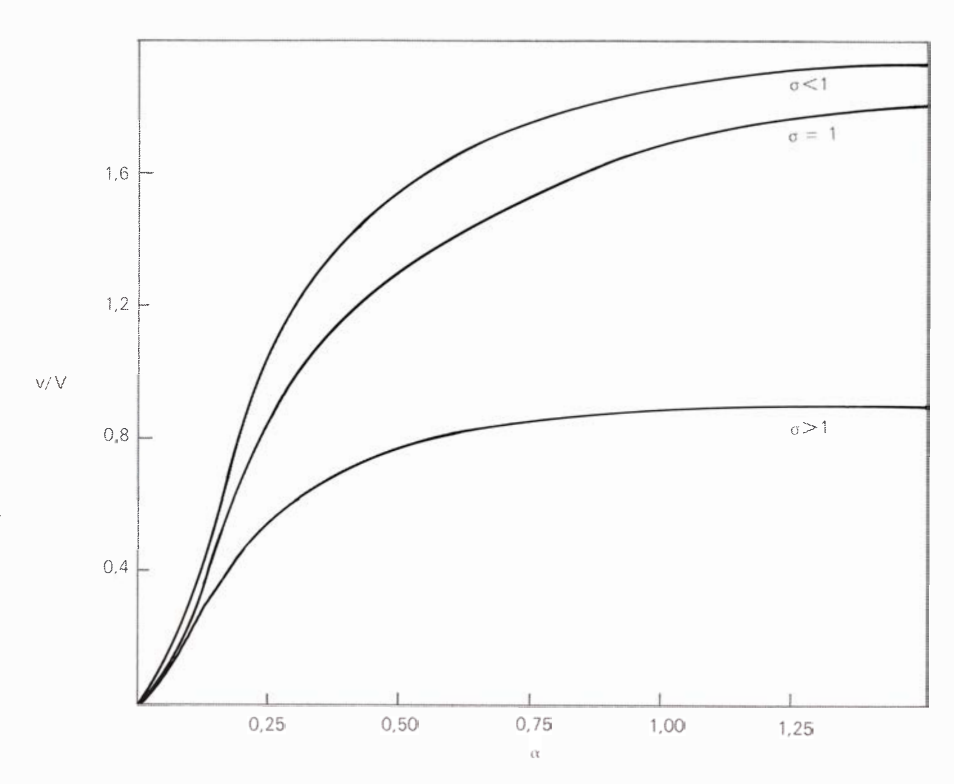
Las expresiones que dan el valor de la proporción de cada forma de la proteína en función de la concentración del ligando se pueden utilizar para calcular el valor promedio del número de centros ocupados por molécula, para cualquier concentración del ligando. La representación directa de la ecuación así obtenida tiene una forma característica que también permite diagnosticar la existencia de cooperatividad y, muy particularmente, si hay cooperatividad positiva en la unión porque, en ese caso, resulta una curva sigmoide.

La saturación de los centros de unión se traduce en una progresiva variación en las propiedades de la proteína. Puede ocurrir que, por cada centro ocupado, se produzca una variación de igual magnitud, tanto si la molécula proteínica tiene el otro centro libre como si lo tiene ocupado. Pero también se puede dar el caso de que esa variación sea mayor o menor, atendiendo precisamente a esa circunstancia. Cuando la eficacia intrínseca de un centro ocupado (contribución de la propiedad por unidad de concentración) depende del estado de los demás centros en la molécula se habla de cooperatividad en la acción y de efecto homotrópico positivo o negativo, según que la ocupación exalte o atenúe la eficacia.

Si la proteína es un enzima con dos centros de unión por molécula para su único sustrato, la eficacia catalítica de un centro puede depender de que el otro esté libre u ocupado. Utilizando las expresiones que permiten calcular la proporción de cada forma en función de la



VALOR PROMEDIO del número de centros ocupados por molécula de un enzima con dos centros equivalentes por molécula para su único sustrato. Dicho valor depende de la concentración de éste. Se puede calcular el valor promedio del número de centros ocupados por molécula, $\bar{X}$, sumando los productos que resultan de multiplicar la proporción de las moléculas del enzima en cada forma por el número de centros ocupados que tiene. La representación de $\bar{X}$, en función de α proporciona una línea cuya forma depende del valor de ϱ. Concretamente, de si existe cooperatividad positiva o negativa o no hay cooperatividad en la unión. Las gráficas de la figura corresponden a $\varrho = 0{,}1$; $\varrho = 1$ y $\varrho = 10$.



PARA OBTENER LA ACTIVIDAD MOLAR DE UN ENZIMA con dos centros de unión equivalentes por molécula para su único sustrato en función de la concentración de éste, basta con sumar los productos que resultan de multiplicar la proporción de moléculas de enzima en cada forma, por la eficacia catalítica de sus centros activos; σ es la relación entre las eficacias catalíticas del enzima a alta y baja concentración de sustrato. La forma de la línea obtenida al representar la actividad molar frente a la concentración de sustrato, α, depende del valor de σ. Es decir de si existe cooperatividad en la catálisis o no y, en caso afirmativo, del signo que tenga. Para las tres últimas figuras se supone que hay cooperatividad positiva en la unión ($\varrho = 10$) y que los valores de sigma son, respectivamente, de 0,9, 1 y 2 (que corresponden a la curva superior, central e inferior). Las ilustraciones son del autor.

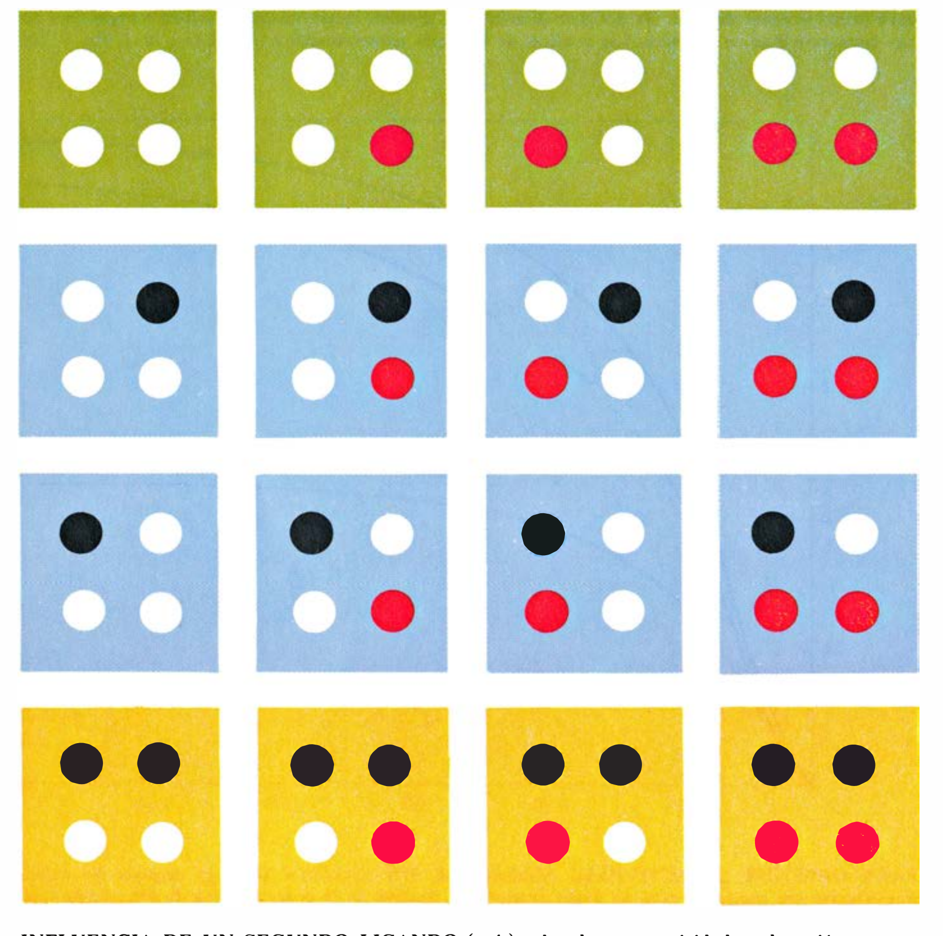
concentración de sustrato y teniendo en cuenta la eficacia intrínseca de cada forma, se puede obtener el valor de la actividad molar del enzima en función de la concentración de sustrato, por medio de una ecuación en la que figuran como parámetros la afinidad por el sustrato y la eficacia catalítica del enzima, a baja concentración de sustrato, la relación entre las afinidades por el sustrato a alta y baja concentración y la relación entre las eficacias en la catálisis del enzima a alta y baja concentración de sustrato. La representación directa de esta función proporciona una curva, cuya forma depende de si existe cooperatividad en la unión y del signo que tenga, así como de si existe cooperatividad en la catálisis y es positiva o negativa [*véanse las ilustraciones de la página anterior*].

Es un hecho comprobado que toda proteína que presenta efectos homotrópicos tiene centros de unión para otros efectores cuya ocupación influye sobre dichos efectos homotrópicos. Esta faceta de las proteínas en general y de los enzimas en particular se describe diciendo que exhiben efectos heterotrópicos. La consecuencia de estos efectos heterotrópicos consiste siempre en una exaltación o disminución de la cooperatividad en la unión y, en su caso, también en la catálisis, que caracteriza al efecto homotrópico. De esta manera, bajo la influencia de un efector se altera el comportamiento del enzima respecto de su sustrato, de forma que aparecen o desaparecen los efectos homotrópicos en la unión y en la catálisis. Sin duda, la cooperatividad en la unión y la catálisis es ventajosa para la célula, pero sólo bajo ciertas circunstancias. Por eso se halla bajo control, de lo contrario constituiría una característica impuesta por la propia naturaleza de la proteína [*véase la ilustración de esta página*].

La representación de Hill aplicada a los datos sobre la actividad de un enzima alostérico permite detectar la existencia de cooperatividad y su signo. Para enzimas con más de dos centros por molécula puede incluso denunciar la presencia de cooperatividad mixta (cooperatividad con diferente signo a medida que progresa la ocupación de los centros). En nuestro Departamento de la Universidad de León, con Dolores de Arriaga, Joaquin Soler, Félix Bustos y Quira Cadenas, estudiante de biología fundamental, trabajando con la deshidrogenasa láctica de micelio del hongo *Phycomyces* hemos encontrado una serie de efectos tanto homotrópicos como heterotrópicos. En particular, es notable el efecto de D-fructosa-1,6-bisfosfato sobre la cooperatividad en la unión con el sustrato y con el coenzima, como sólo puede hacerlo un efector alostérico. Por un lado, reduce la cooperatividad positiva para NADH y, por otro, transforma en mixta la cooperatividad para el piruvato [*véase la ilustración de la página siguiente*].

La existencia de cooperatividad en la unión parece implicar algún tipo de conexión entre los centros de unión en una molécula. En un principio se supuso que la conexión ocurriría de forma directa y así se imaginó que los grupos hemo se hallarían concentrados en algún área de la arquitectura proteínica de la hemoglobina. No obstante, los resultados de difracción de rayos X obtenidos por M. Perutz invalidaron tal presunción al revelar que los grupos hemo se hallan hundidos en unos bolsillos, cuyos bordes hacen imposible cualquier comunicación. En tales circunstancias se hizo preciso un planteamiento totalmente distinto. Cuando Monod, Wyman y Changeux lo proponen en 1965 le dan el carácter de un modelo, en que los efectos homotrópicos y heterotrópicos de las proteínas alostéricas encuentran explicación a través de interacciones en las que no participan directamente los centros de unión.

Un elemento muy obvio del modelo alostérico es la suposición de que la proteína reguladora debe estar constituida por la agregación de subunidades iguales, que se denominan protómeros, que ocupan posiciones equivalentes en el agregado, que se denomina oligómero. Ya se conocía este tipo de organización en la hemoglobina, pero el modelo la propone como base de la cooperatividad y, por tanto, supone que la comparten todos los enzimas reguladores. Con independencia de sus méritos propios, esta tesis ha despertado un inusitado interés por los estudios sobre la estructura de las proteínas que desempeñan un papel clave en distintas áreas de la regulación biológica, no sólo los enzimas reguladores, sino también entidades como los represores del modelo del operón y los receptores para las hormonas, tanto citoplasmáticos como situados en la membrana celular. Como consecuencia de estos estudios se ha logrado con-

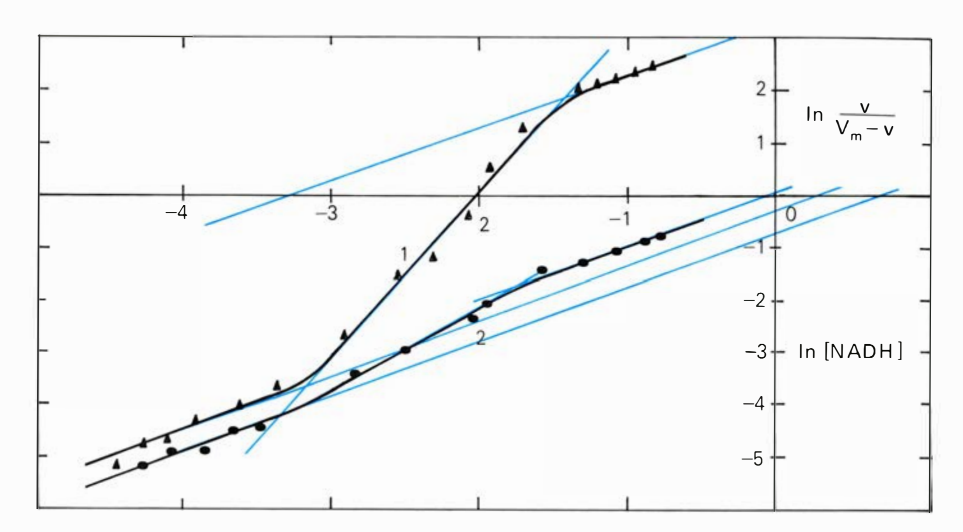


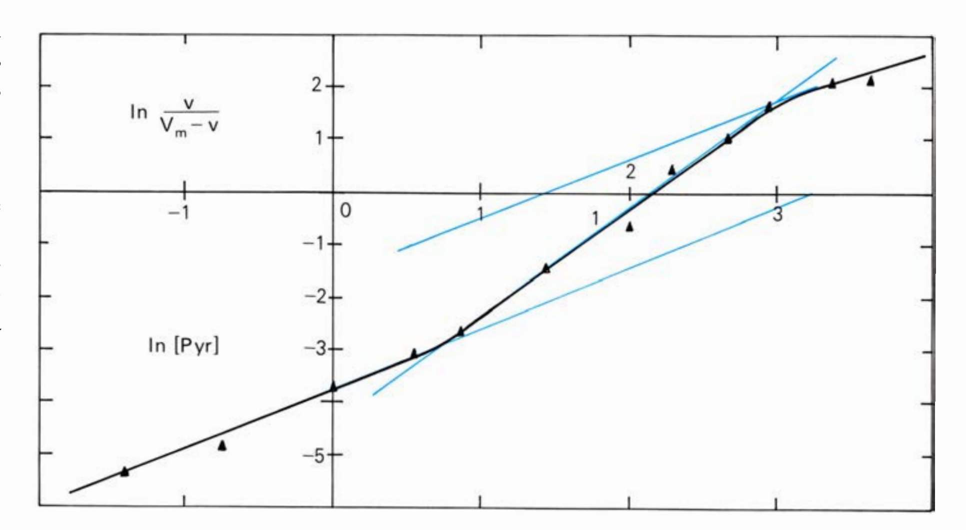
INFLUENCIA DE UN SEGUNDO LIGANDO (*rojo*) sobre la cooperatividad en la unión con un primer ligando (*negro*). Si se trata de un enzima con dos centros de unión equivalentes por molécula para su único sustrato, el segundo ligando puede ser un efector que influye sobre la unión con el sustrato. Un inventario de todas las formas de la proteína que existen en presencia de uno y otro ligando como éste esquematizado, sirve de base para desarrollar las ecuaciones que describen la influencia del efector sobre la cooperatividad en la unión con el sustrato y la cooperatividad en la catálisis.

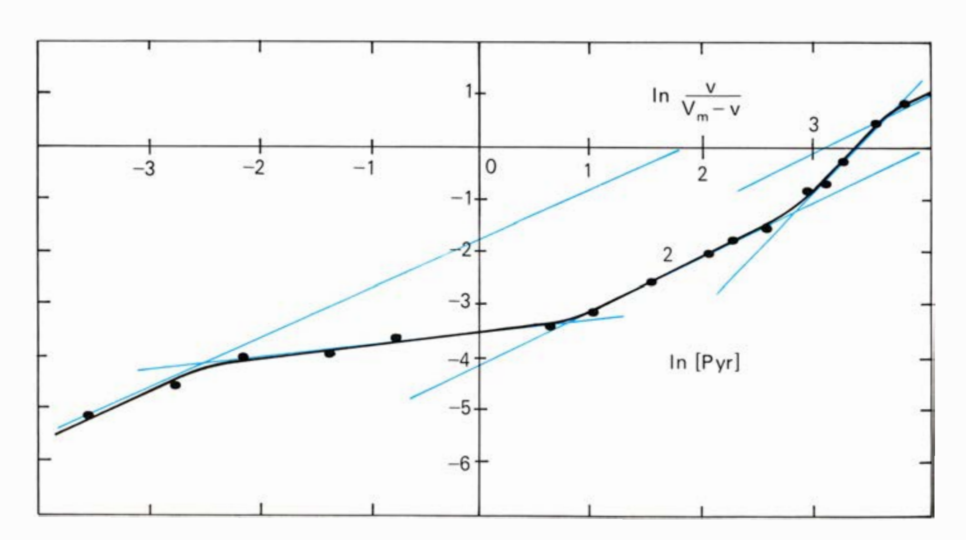
firmar la universalidad de la organización oligomérica e ilustrar algunas de las variadas potencialidades que encierra. La arquitectura de la molécula proteínica se sustenta sobre interacciones entre sus aminoácidos constitutivos, pero también sobre estas interacciones descansan sus variadas funciones biológicas. En la estructura de la proteína se pueden distinguir varios niveles de organización. Las interacciones a cada nivel generan curiosas propiedades de la proteína y son precisamente las interacciones al nivel del oligómero las responsables de la cooperatividad que presentan las proteínas alostéricas.

Hoy poseemos una noción, al menos aproximada, del proceso por el cual se originan las moléculas de una proteína. El progreso se ha logrado por dos caminos diferentes. Por un lado, están los estudios sobre la maquinaria enzimática en que unos 200 componentes distintos intervienen en la polimerización de los veinte aminoácidos característicos de las proteínas, para generar la estructura primaria propia de cada proteína. Por otro, las investigaciones sobre la anatomía de la molécula proteínica, que permiten inferir el programa que sigue el plegamiento de la cadena polipeptídica, a medida que se va formando según progresa la polimerización.

Al mismo tiempo que crece la proteína por incorporación de sucesivos aminoácidos se multiplican las posibilidades que los mismos tienen de participar en interacciones intramoleculares. En primer término, están aquellas en que intervienen los átomos de oxígeno y nitrógeno directamente vinculados a los que forman los propios enlaces peptídicos que mantienen los aminoácidos engarzados en la cadena conforme a un orden o secuencia específicos. Estas interacciones se materializan en la formación de enlaces de hidrógeno que sirven de apoyo a la llamada estructura secundaria de la proteína. Elementos de esta estructura son la hélice α y las hojas corrugadas β. Muy probablemente su generación tentativa viene a servir de germen a la estructura definitiva de la proteína. Después surgen interacciones en que participan las cadenas laterales de los residuos de los aminoácidos (llamados también restos aminoácidos). Por ello dependen muy específicamente del orden o secuencia en que se encuentran dispuestos a lo largo de la cadena polipeptídica. Estas interacciones establecen dónde se formará definitivamente una estructura secundaria determinada y qué residuos de aminoácido se verán implicados en ella.







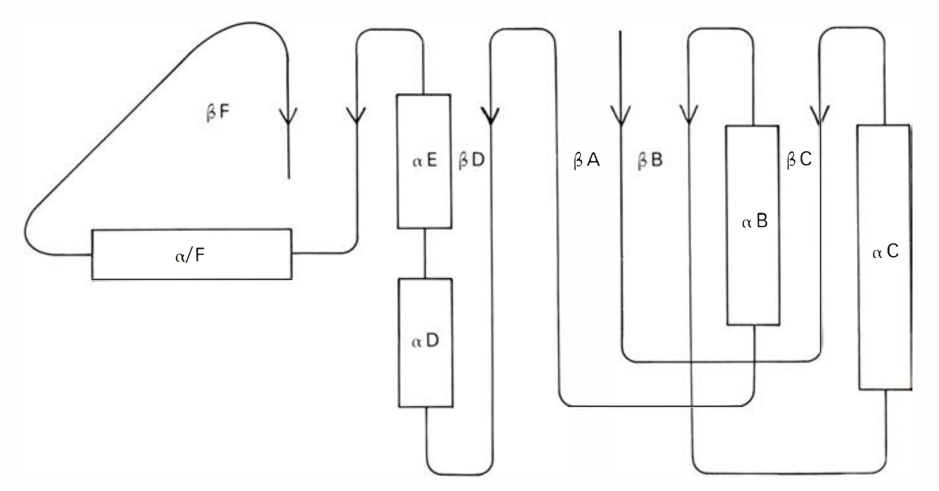
INFLUENCIA DEL EFECTOR D-FRUCTOSA-1,6-BIS FOSFATO sobre la cooperatividad para el sustrato (piruvato, pyr) y la coenzima (NADH) de la deshidrogenasa láctica de *Phycomyces blakesleeanus* a *p*H 7,7. Las líneas 1 y 2 en la representación de Hill (*ilustración superior*) corresponden a los datos obtenidos en ausencia y presencia de efector que provoca una marcada reducción en el grado de cooperatividad positiva para el coenzima en cuestión. La comparación de las líneas 1 (*ilustración central*) y 2 (*ilustración inferior*) en ausencia y presencia de efector permiten detectar la transformación de la cooperatividad pura en mixta por lo que se refiere al sustrato. Evidentemente, se trata de un efector alostérico.

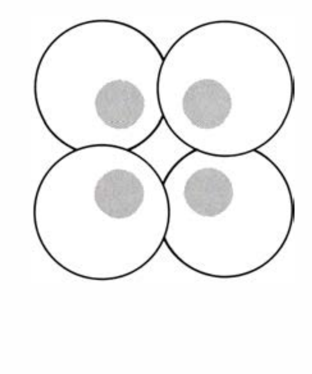
Además las interacciones entre los aminoácidos dirigen el plegamiento sobre sí mismos de los diversos tramos de estructura secundaria hasta generar regiones autónomas de la molécula proteínica dotadas de una característica estructura supersecundaria. Estas regiones que vienen a ser como proteínas globulares en miniatura, se construyen de acuerdo con principios generales de montaje, como el que establece que deben tener un núcleo hidrofóbico mientras los residuos polares deben estar expuestos al disolvente acuoso. Pero también se organizan conforme a requisitos especiales, ya que en cada una de estas regiones, que se denominan dominios o centros de unión, tiene su sede la capacidad de unión específica con determinado ligando.

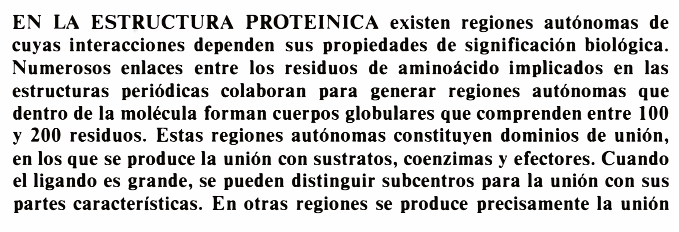
Dependiendo del tamaño de la cadena polipeptídica, en una molécula proteínica pueden coexistir mayor o menor número de regiones autónomas y así dotarla de mayor o menor número de funciones. La forma como se disponen estas regiones unas respecto de otras viene a constituir otro nivel de organización, la estructura terciaria. A este nivel se observan interacciones entre dominios de las que depende la actividad biológica propia de la proteína, como ocurre en la interacción de los centros de unión para los sustratos de un enzima de la que depende su eficacia catalítica. Sin duda, un caso extremo en este aspecto lo constituyen las proteínas multifuncionales, que presentan numerosos dominios de unión y variadas actividades enzimáticas con sede en una misma cadena polipeptídica. En la mayoría de los casos, sin embargo, se logra acumular distintas funciones en la misma molécula proteínica utilizando un recurso distinto. Se produce en la célula la agregación espontánea y altamente específica de varias cadenas polipeptídicas, cada una de las cuales aporta sus dominios propios. Entre estos dominios están precisamente las áreas de unión específicas que permiten el ensamblaje de las subunidades y de cuyas interacciones depende la llamada estructura cuaternaria de la proteína [*véase la ilustración inferior*].

Como se ha descrito, la propia estructura de la molécula proteínica sugiere el programa que sigue su edificación en la célula pero, además, da pie a especulaciones sobre las ventajas que le puede proporcionar tanto en lo referente a simplificar su tarea, como lo que respecta a las cualidades que posee el producto acabado. La organización cuaternaria hace posible la obtención de moléculas de gran tamaño dotadas de una variedad de funciones, a partir de cadenas polipeptídicas de longitud razonable. Se ha sostenido que la conveniencia de este diseño estriba en que un error cometido en la fase de polimerización sólo afectaría a una parte de la molécula y no determinaría el rechazo de toda ella. La agregación, intercalada en el proceso de construcción, vendría a constituir una etapa de control de calidad para la proteína. Sin embargo, esta razón no parece decisiva, puesto que existen numerosas proteínas multifuncionales, cuya molécula es una cadena polipeptídica ininterrumpida a pesar de su extensión. Parece más probable que la trascendencia de la estructura cuaternaria derive de que hace posible la estructura oligomérica. De hecho, la gran mayoría de las proteínas con estructura cuaternaria son oligómeros.

En el modelo que Monod, Wyman y Changeux propusieron en 1965 se admite que la proteína es un oligómero que se puede presentar en dos conformaciones diferentes relativamente estables. Los protómeros que constituyen el oligómero deben tener la misma conformación, pero esta conformación será diferente en uno y otro confórmero del oligómero. Se supone también que los dos confórmeros del oligómero se encuentran permanentemente en equilibrio a través de una reacción en la que se produce la transición de una a otra conformación (transición alostérica). La proteína tiene varios centros de unión para un determinado ligando, tantos como protómeros posea el oligómero, y todos los centros de unión de un confórmero son equivalentes y poseen la misma afinidad por el ligando, sin que influya en ella la ocupación de los centros en la misma molécula. Con anterioridad (1964), Wyman ya había señalado que la constante de equilibrio para una reacción entre moléculas que poseen centros de unión para un mismo ligando tiene un valor que depende de la concentración del ligando siempre y cuando las moléculas difieran en su afinidad por el ligando. Pero, al proponer el modelo alostérico, que luego se calificará de concertado, se insiste sobre el efecto recíproco que sirve de base a la cooperatividad. Aunque esta propiedad no existe en la unión del ligando con uno y otro de los confórmeros, su unión con ambos en equilibrio a través de la reacción de transición alostérica puede mostrar coo-







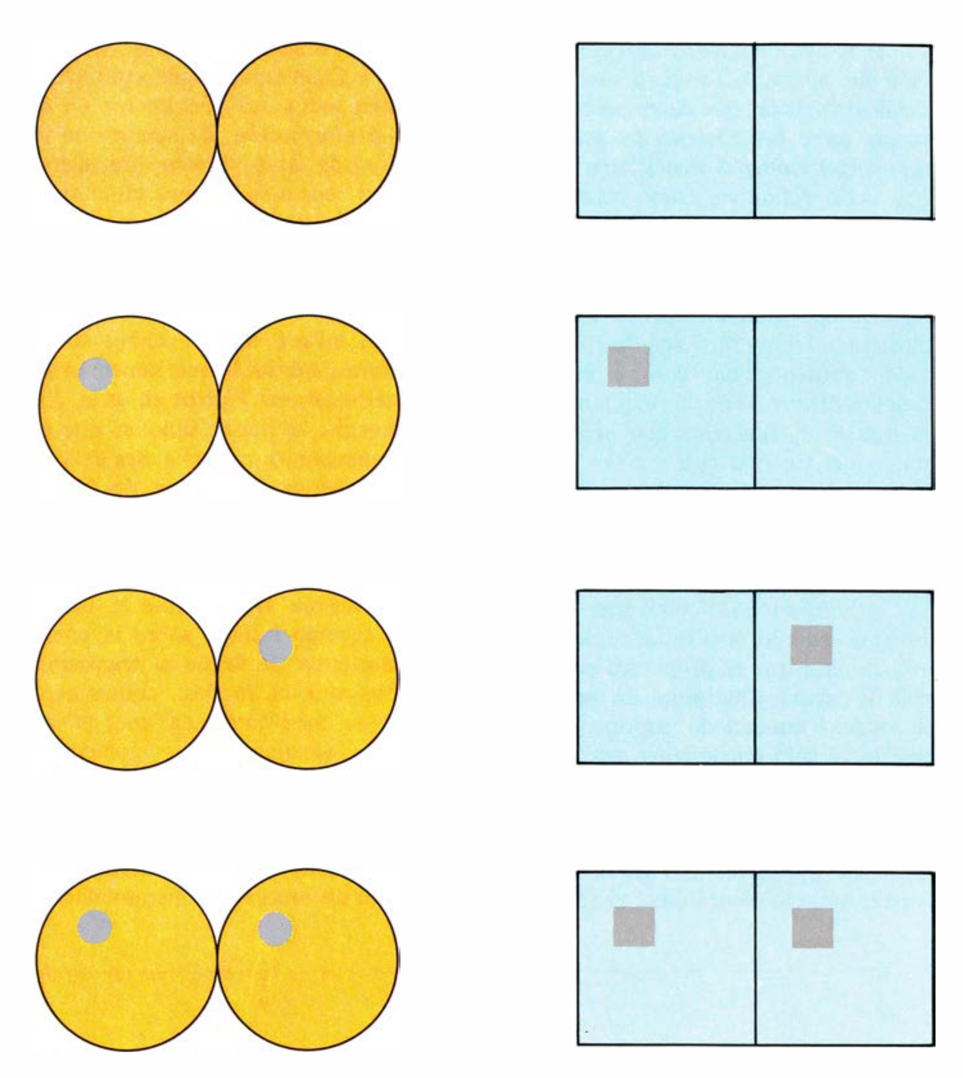
EN LA ESTRUCTURA PROTEINICA existen regiones autónomas de cuyas interacciones dependen sus propiedades de significación biológica. Numerosos enlaces entre los residuos de aminoácido implicados en las estructuras periódicas colaboran para generar regiones autónomas que dentro de la molécula forman cuerpos globulares que comprenden entre 100 y 200 residuos. Estas regiones autónomas constituyen dominios de unión, en los que se produce la unión con sustratos, coenzimas y efectores. Cuando el ligando es grande, se pueden distinguir subcentros para la unión con sus partes características. En otras regiones se produce precisamente la unión específica entre las subunidades para constituir el agregado. De la interacción de los dominios depende tanto la catálisis enzimática, como su regulación por efectores. A la izquierda, se representa el dominio de unión con NAD en cada una de las subunidades de la molécula de malato deshidrogenasa soluble (los rectángulos simbolizan hélices α, las rectas paralelas verticales representan los seis tramos de hoja paralela β típicos de este dominio en las deshidrogenasas). A la derecha, un oligómero. Se trata del tetrámero *D*-gliceraldehído-3-fosfato deshidrogenasa de langosta. (En cada protómero se señala la posición del dominio de unión con NAD.)

peratividad, si los dos confórmeros poseen distinta afinidad por el ligando. Analíticamente resulta bastante sencillo comprobar este hecho, que bien se podría describir como "efecto alostérico".

Como ya se adelantó, el origen de la cooperatividad en la unión que exhibe el modelo en su comportamiento no hay que buscarlo en un mecanismo por el cual la estructura de la proteína se modifique al unirse al ligando. Radica, en cambio, en un quimismo; se basa en las curiosas propiedades que caracterizan a las reacciones químicas acopladas. Cuando la unión con el ligando modifica el valor de la constante de equilibrio para la transición de una a otra conformación, favorece precisamente el aumento en la concentración del confórmero de mayor afinidad por el ligando. De esta manera, crece la proporción de centros situados en moléculas con la conformación de mayor afinidad, lo que evidentemente equivale a un aumento progresivo de la afinidad por el ligando a medida que se van ocupando los centros de unión. De ahí la aparente cooperatividad en la unión que se observa en el sistema, sin que exista en sus partes.

Con el fin de contar con una descripción cuantitativa del origen de la cooperatividad en el modelo concertado, hemos realizado un análisis de la expresión que proporciona el valor de la relación entre las afinidades observadas a alta y baja concentración del ligando, porque esta relación viene a determinar el grado de cooperatividad que presenta el sistema. La clave del análisis radica en la utilización de la relación de afinidades por el ligando de uno y otro confórmero como unidad natural para expresar el valor de la constante de equilibrio para la transición alostérica (constante alostérica).

Del resultado del análisis aplicado a un dímero se concluye que no existe cooperatividad cuando cualquiera de los dos confórmeros está exageradamente favorecido en el equilibrio de la transición alostérica. También se prueba que la cooperatividad es máxima cuando la constante alostérica tiene el mismo valor que la relación entre las afinidades por el ligando (sin que importe cuál de las dos se tome como referencia). En otras palabras, la cooperatividad en la unión es máxima para el dímero, si la energía libre normalizada para la reacción de transición alostérica es igual a la diferencia entre las energías de unión del ligando con uno y otro confórmero. Cualquier desviación de esta coincidencia conduce necesariamente a una disminución de la cooperatividad en la unión, que, por supuesto, sólo puede ser positiva (o nula) [*véase la ilustración de esta página*]. La aplicación del mismo tipo de análisis a oligómeros con diferentes números de protómeros conduce también a conclusiones interesantes en las que, sin embargo, no entraremos, salvo para indicar que el valor máximo de la cooperatividad es más pronunciado cuanto mayor sea la desproporción entre las afinidades de los dos confórmeros por el ligando.



EL ENZIMA ES UN DIMERO que se comporta conforme al modelo alostérico concertado. En ausencia de ligandos, las únicas formas del enzima son las que aparecen en la primera fila de la tabla. Se trata de dos confórmeros en equilibrio a través de una reacción química que se denomina transición alostérica y cuya constante de equilibrio se representa por *L* (se considera el confórmero simbolizado por dos círculos, como reaccionante). En presencia de sustrato existe un mayor número de formas de enzima que difieren entre sí por el grado de ocupación que posean de los centros específicos de unión.

El modelo alostérico concertado permite explicar la cooperatividad positiva en la unión del oxígeno con la hemoglobina y con la misma facilidad da cuenta de la cooperatividad en la unión de un enzima alostérico con sus sustratos. Pero en este último caso, los supuestos básicos del modelo se pueden utilizar como lo hemos hecho nosotros, para justificar también la existencia de cooperatividad en la catálisis y establecer las condiciones que se deben cumplir para que sea positiva o negativa. Ya hemos visto que para que haya cooperatividad en la unión, los dos confórmeros del oligómero deben poseer distinta afinidad por el ligando. Para que un enzima alostérico presente cooperatividad en la catálisis, sus confórmeros deben tener no sólo distinta afinidad por el sustrato sino también distinta eficacia catalítica. Por tanto, no puede haber cooperatividad en la catálisis si no hay cooperatividad en la unión, aunque lo contrario sí puede ocurrir. La cooperatividad en la catálisis será positiva, si un confórmero aventaja al otro, tanto en afinidad como en eficacia catalítica. Por el contrario, cuando un confórmero aventaje al otro solamente en uno de los aspectos señalados, la cooperatividad en la catálisis será negativa. Estas conclusiones se ilustran para el dímero concertado, pero es interesante señalar que tienen validez para cualquier oligómero. En el caso del dímero se ha visto la influencia que ejercen diversos factores sobre la relación entre las eficacias catalíticas a alta y baja concentración de sustrato (sobre la ener-

gía de activación a alta y baja concentración de sustrato). En particular, se ha detallado el papel que desempeña la diferencia entre las energías de activación que corresponden a uno y otro confórmero. En definitiva, como ocurre para la cooperatividad en la unión, la cooperatividad en la catálisis que no existe para ninguno de los dos confórmeros, hace su aparición a través de un efecto alostérico. Uno y otro tipo de cooperatividad surgen como consecuencia del desplazamiento en el equilibrio de la transición alostérica que provoca la unión del sustrato con los dos confórmeros del enzima [*véase la ilustración de la página siguiente*].

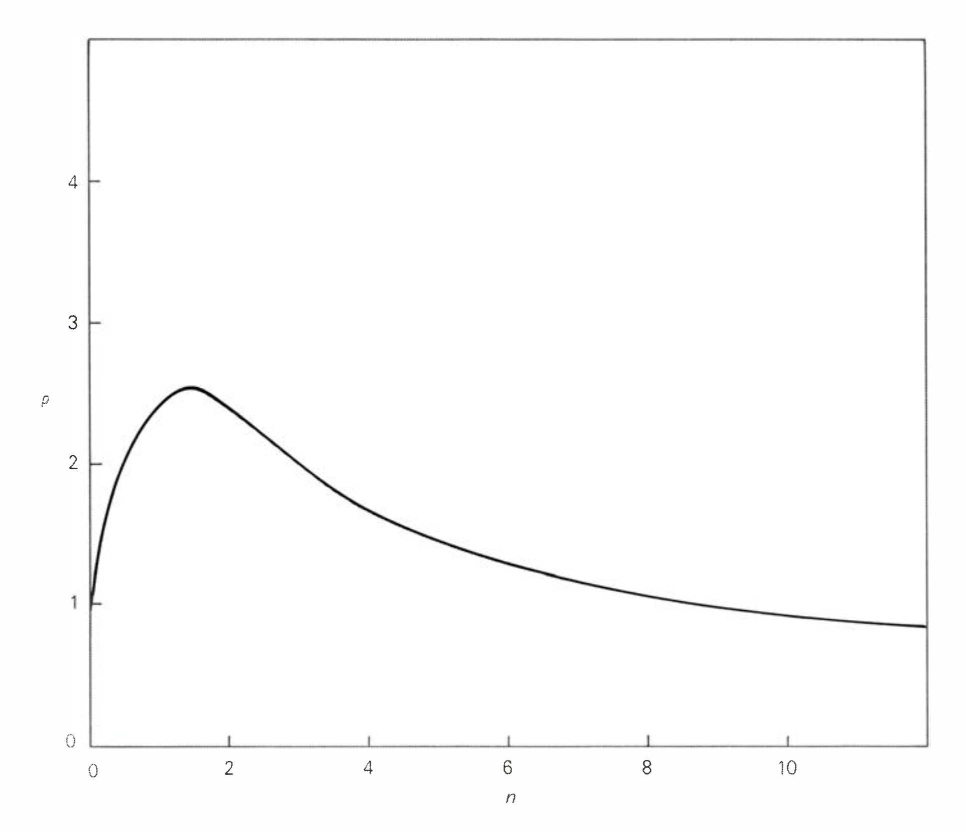
Con las ideas de Koshland, Nemethy y Filmer (1966) como base ha ido tomando cuerpo otro modelo alostérico más general que el propuesto por Monod, Wyman y Changeux. En tanto que el modelo concertado supone que el oligómero sólo puede tener dos confórmeros, sin que importe el número de protómeros que comprenda, el modelo generalizado implica un número de confórmeros que crece a medida que lo hace el número de protómeros que comprende el oligomero. Para Monod, Wyman y Changeux, los dos confórmeros tienen todos sus protómeros en la misma conformación, de suerte que la transición de uno al otro (transición alostérica) ocurre de manera simultánea (concertada) para todos los protómeros. El modelo generalizado admite, en cambio, que pueden existir tanto confórmeros puros (que tienen todos los protómeros en la misma conformación) como confórmeros híbridos (que tienen unos protómeros en una y otros en otra). En consecuencia, se puede suponer que el paso de una conformación a otra del oligómero generalizado ocurre de forma escalonada (secuencial) por transición de la conformación de uno solo de los protómetros cada vez.

De la misma manera que la unión con un ligando influye sobre la constante alostérica de la única transición que considera el modelo concertado, también lo hace con cada una de las constantes alostéricas que gobiernan cada una de las transiciones escalonadas del modelo generalizado. También ocurre que la unión del ligando con cada uno de los confórmeros del modelo generalizado no muestra cooperatividad y, no obstante, la unión con todos ellos en equilibrio a través de las diversas transiciones alostéricas exhibe cooperatividad. El modelo generalizado presenta una característica que no tenía el concertado; sin embargo, esta característica es la que le confiere una clara ventaja sobre aquél. El modelo generalizado puede explicar la existencia de cooperatividad en la unión tanto positiva como negativa.

La cooperatividad negativa en la unión pasó desapercibida sin duda porque a diferencia de lo que ocurre con la cooperatividad positiva, la representación directa de la ecuación de velocidad de un enzima que posea esta propiedad parece proporcionar una hipérbola. No obstante, ya en el momento de la formulación del modelo de Koshland, Nemethy y Filmer se señaló la existencia de este tipo de comportamiento en cierto número de enzimas. Con el estímulo que representó la aparición de este modelo y con la utilización de otras representaciones gráficas de la ecuación de velocidad que, como las basadas en las que en su día propusieron Hill y Scatchard, permiten diagnosticar con la misma facilidad la cooperatividad positiva que la negativa, se inició una búsqueda de ambos tipos de comportamiento en todos los enzimas reguladores. Como resultado se ha establecido que los dos tipos de comportamiento deben tener gran importancia en la regulación. Es probable que el significado metabólico de la cooperatividad negativa en la unión haya que buscarlo en la garantía que proporciona de una marcada permanencia del sustrato o coenzima en los centros que ocupa cuando su concentración disminuye. De esta manera, el enzima se vuelve insensible en cierta medida a la variación en la concentración de un sustrato o coenzima, permaneciendo listo para actuar tan pronto como disponga de una concentración adecuada del otro sustrato.

Partiendo de los supuestos básicos del modelo generalizado se puede realizar un análisis de la influencia de diversos factores sobre su comportamiento. Como en el caso del modelo concertado, el objeto del análisis será doble. Por un lado, permitirá profundizar en el origen del efecto alostérico en términos de un quimismo. Por otro, al describir el efecto de factores como el valor de las constantes alostéricas, hará posible la comprensión de la regulación por un efector del comportamiento del enzima alostérico, puesto que el efector precisamente modifica el valor de algunos de esos factores.

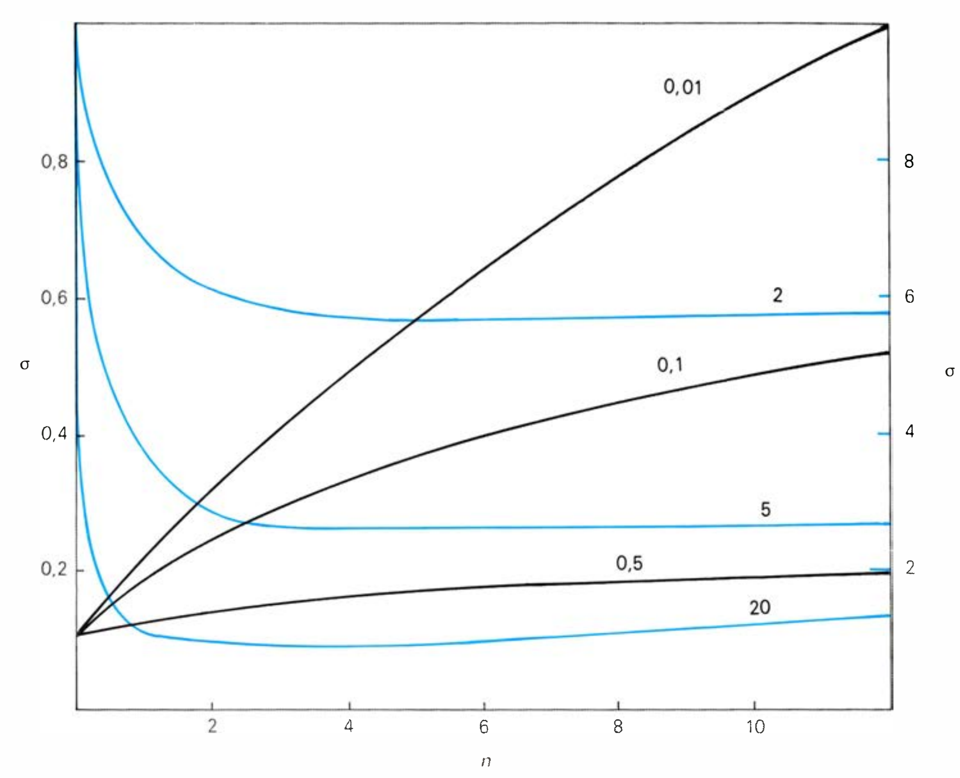
Para realizar el análisis con un enzima que se comporte conforme a los supuestos del dímero generalizado, al objeto de poner de manifiesto los factores de que



INFLUENCIA DEL VALOR DE LA CONSTANTE ALOSTERICA sobre el grado de cooperatividad en la unión con su único sustrato para un enzima que es un dímero concertado. Para poder apreciar la influencia del valor de la constante alostérica, *L*, sobre la cooperatividad en la unión con el sustrato se hacen dos sustituciones. Por un lado, se hace desaparecer de la expresión la concentración de sustrato dada en dos unidades distintas y se sustituye por la relación entre sus afinidades por uno y otro confórmero. Por otro lado, el valor de *L* se expresa por *n* tomando como unidad "natural" el valor de *c*. De esta forma, el valor de ϱ resulta ser una curiosa función del valor *n*. La función vale igual para *n* que para su inverso, sea cual fuere el valor de *c*. Y vale lo mismo para cualquier valor de *n*, si *c* tiene determinado valor o su inverso. La cooperatividad tiende a desaparecer si *L* es muy grande o muy pequeña y tiene un valor máximo para *n* = 1 para todo *c*. La gráfica se ha hecho para *c* = 100, o *c* = 0,01.

depende la cooperatividad en la unión, hemos introducido el uso de algunas relaciones entre factores, relaciones que en el fondo corresponden a diferencias de energía entre dos reacciones de las que el modelo comprende. Hemos utilizado como variables: la relación de afinidades por el sustrato de una y otra conformación del protómero, la relación entre los valores de la segunda y la primera constante alostérica y se ha referido el valor de la primera constante alostérica al valor de la relación de afinidades, para así expresar el valor de L en unidades "naturales". De este modo, disponiendo de una ecuación apropiada, se ha podido calcular el valor de la relación entre la afinidad por el sustrato a alta y baja concentración (como medida del grado de cooperatividad en la unión) en función de las variables indicadas [*véase la ilustración de la página siguiente*].

Las conclusiones del análisis, que resultan muy evidentes mediante la utilización de representaciones gráficas, abren efectivamente el camino hacia la comprensión de la biología molecular del efecto alostérico. En primer término, se comprueba que la relación entre los valores de las constantes alostéricas es el factor determinante del signo de la cooperatividad en la unión. Precisamente la cooperatividad será positiva o negativa según que la relación entre la segunda y la primera constante alostérica sea superior o inferior a la unidad. Esto quiere decir que la cooperatividad en la unión será positiva si las formas puras están favorecidas respecto de las formas híbridas, y negativa si ocurre lo contrario. Se puede decir que este resultado concuerda con el obtenido para el modelo concertado, si se piensa que en aquél se admite una preferencia absoluta por las formas puras, lo que supone que sólo pueda explicar la cooperatividad positiva. También el valor de la relación entre las dos constantes alostéricas afecta al grado de cooperatividad, siendo el efecto tanto más pronunciado cuanto mayor sea la relación, si la cooperatividad es positiva, y cuanto menor sea, si la cooperatividad es negativa. En definitiva, el efecto es tanto más intenso cuanto mayor sea la preferencia que exista entre los confórmeros puros y los mixtos, sea cual fuera el sentido en que se establezca. Además se puede notar que el valor de la primera constante alostérica, referido al valor de la relación de afinidades, influye por sí mismo en el grado de cooperatividad. En particular, la cooperatividad desaparece para valores muy altos o muy bajos de la constante alostérica y presenta un valor óptimo, a cuyos dos lados el grado de cooperatividad varía con ritmos muy diferentes. Por último, conviene señalar que ningún factor puede ejercer su influencia sobre la cooperatividad a menos que las dos conformaciones del protómero posean distinta afinidad por el sustrato, de modo que este factor tiene un carácter permisivo respecto de los demás, como ocurre también para el modelo concertado.



INFLUENCIA DEL VALOR DE LA CONSTANTE ALOSTERICA sobre la cooperatividad en la catálisis para un enzima que es un dímero concertado. Para poder apreciar la influencia del valor de la constante alostérica, *L*, sobre la cooperatividad en la catálisis se hacen las mismas sustituciones que en el caso de la cooperatividad en la unión y además se introduce el valor de la relación entre las eficacias catalíticas de los dos confórmeros, γ. De este modo el valor de σ resulta ser función del valor de *n*. En azul se puede ver que para $\gamma > 1$ la cooperatividad en la catálisis es negativa (*escala de la izquierda*) y el efecto tanto más pronunciado cuanto mayor es el valor de γ. En negro se puede apreciar que para $\gamma < 1$ la cooperatividad en la catálisis es positiva (*escala de la derecha*) y el efecto tanto más pronunciado cuanto menor es el valor de γ. (Los números superpuestos sobre cada gráfica indican el valor de γ). Para todas las gráficas recogidas en la ilustración se ha considerado que el valor de $c = 100$.

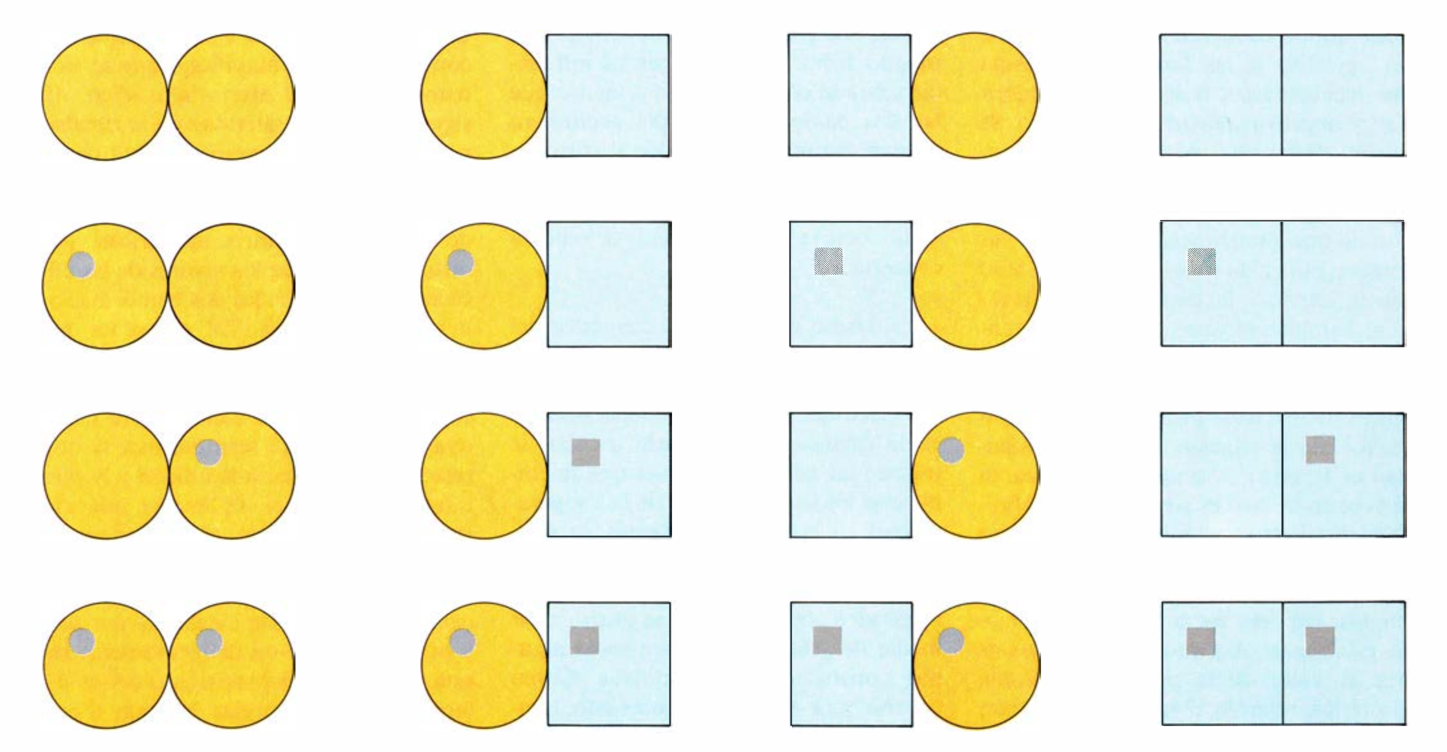
Asimismo en base a los supuestos del modelo generalizado se puede realizar un análisis de la influencia de diversos factores sobre la cooperatividad en la catálisis. Para llevarlo a cabo, se utilizan las mismas variables que se emplearon en la descripción de la cooperatividad en la unión y, además, la relación entre las eficacias catalíticas de las dos conformaciones del protómero. Los resultados obtenidos, que se ilustran por medio de gráficas, permiten sacar algunas conclusiones significativas. Como ocurría para el modelo concertado, tampoco puede haber cooperatividad en la catálisis, si no hay cooperatividad en la unión o si no existe una diferencia en las eficacias catalíticas de las dos conformaciones del protómero. De esta manera, y para el modelo generalizado como para el concertado, se señala un doble origen de la cooperatividad en la catálisis, así como se pone de manifiesto que se fundamenta en un efecto alostérico. El signo de la cooperatividad en la catálisis para el modelo generalizado está determinado conjuntamente por dos factores. Por un lado, la relación entre los valores de las dos constantes alostéricas; por otro, la relación de los valores de las eficacias catalíticas de las dos conformaciones del protómero. Tal como las hemos definido, si ambas relaciones son superiores o inferiores a la unidad, la cooperatividad en la catálisis será negativa, mientras que será positiva si una relación es superior a la unidad y la otra inferior. Por tanto, se detecta una curiosa conexión entre la diferencia de energía para las dos reacciones de transición y la diferencia de energía de activación para la reacción catalizada por una y otra conformación del protómero. De esta conexión entre los signos de las diferencias en las energías depende el que se observe macroscópicamente uno u otro tipo de cooperatividad y que la célula disponga de las ventajas respectivas que le puedan aportar. Una de las diferencias depende sólo de la naturaleza de la proteína, o mejor, de las interacciones

entre sus protómeros en una u otra conformación. La otra diferencia debe depender tanto de la naturaleza de la proteína, o mejor, de la índole de las conformaciones de sus protómeros, como de la naturaleza del sustrato ya que corresponde a las características de los estados de transición con una y otra conformación del protómero [*véase la figura de la página 118*].

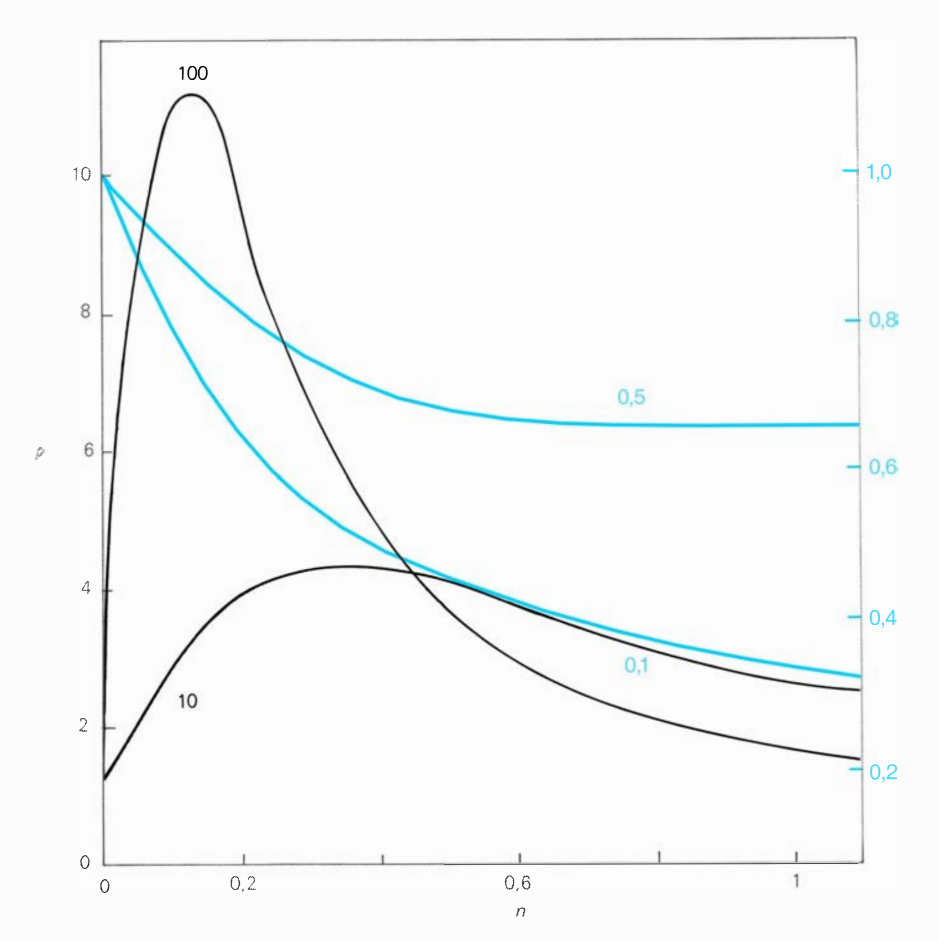
Como se puede apreciar en las ilustraciones, la cooperatividad es tanto más intensa cuanto mayor sea la diferencia en las energías de activación. También se puede comprobar que el propio valor de la primera constante alostérica (referido a la relación entre las afinidades por el sustrato de una y otra conformación del protómero) influye sobre el grado de cooperatividad. Como en todos los casos que hemos mencionado anteriormente, la cooperatividad desaparece para valores muy altos o muy bajos de la constante alostérica y tiene un valor óptimo, a cuyos dos lados el grado de cooperatividad disminuye a ritmos muy diferentes.

Sin duda la clave de la biología molecular de la regulación reside en la forma en que un efector influye sobre el comportamiento del enzima alostérico. Por ello bien merece la pena hacer unas consideraciones aunque sean someras sobre cómo actúa el efector dentro del marco de los modelos alostéricos descritos. La influencia del efector sobre la cooperatividad en la unión no tiene mayor complicación. Como ya hemos señalado repetidas veces, la unión con el efector desplaza el equilibrio entre las diversas conformaciones del enzima y de este modo altera el grado de cooperatividad en la unión con el sustrato. La influencia del efector sobre la cooperatividad en la catálisis es más compleja. Para dar cuenta de la influencia de un efector sobre la cooperatividad en la unión basta con conocer su afinidad por una y otra conformación del protómero. En cambio, para determinar cómo afecta su unión con el enzima la cooperatividad en la catálisis es preciso que conozcamos además cuál es la eficacia catalítica de una y otra conformación del protómero, cuando su centro para el efector está ocupado. Es decir, cómo se altera la eficacia catalítica del protómero en sus dos conformaciones, cuando se une al efector. Este tipo de influencia es terciaria, es decir, se ejerce a nivel del protómero y no depende de la influencia cuaternaria o global, que se manifiesta gracias al efecto alostérico. Según esto, la unión con un efector puede tener dos efectos bien diferenciados sobre un enzima. Por un lado, puede desplazar el equilibrio preexistente entre sus confórmeros bien determinados; por otro, puede alterar la eficacia catalítica del centro activo en el protómero con que se encuentre unido, quizás a través de un cambio de conformación inducido en la región de ese centro activo, como puede ocurrir en un enzima ordinario. Según sus afinidades respectivas por los confórmeros del enzima, dos efectores producirán desplazamientos diferentes en su equilibrio y conducirán a variaciones diferentes de la cooperatividad en la unión pero no serán capaces de alterar la afinidad por el sustrato que caracteriza a cada una de las dos conformaciones del protómero. En cambio, será muy posible que uno y otro efector hagan variar la eficacia catalítica de las dos conformaciones globales del protómero de maneras diferentes.

Numerosos experimentos de diversos grupos de investigadores, y en particular los realizados por el grupo de Koshland, han puesto de manifiesto que la unión de un ligando con un enzima provoca cambios en las propiedades del enzima que varían según cuál sea la naturaleza del ligando y el tipo de propiedad que se estudie. Tomados en su conjunto vienen a señalar que en la estructura proteica se pueden producir cambios de conformación de muy variada índole, que deben estar netamente localizados y ser autónomos y que afectan el comportamiento de la proteína de muy diversas maneras. Estos resultados se pueden interpretar como hace Koshland, en base a su fructífera idea del ajuste inducido, que efec-



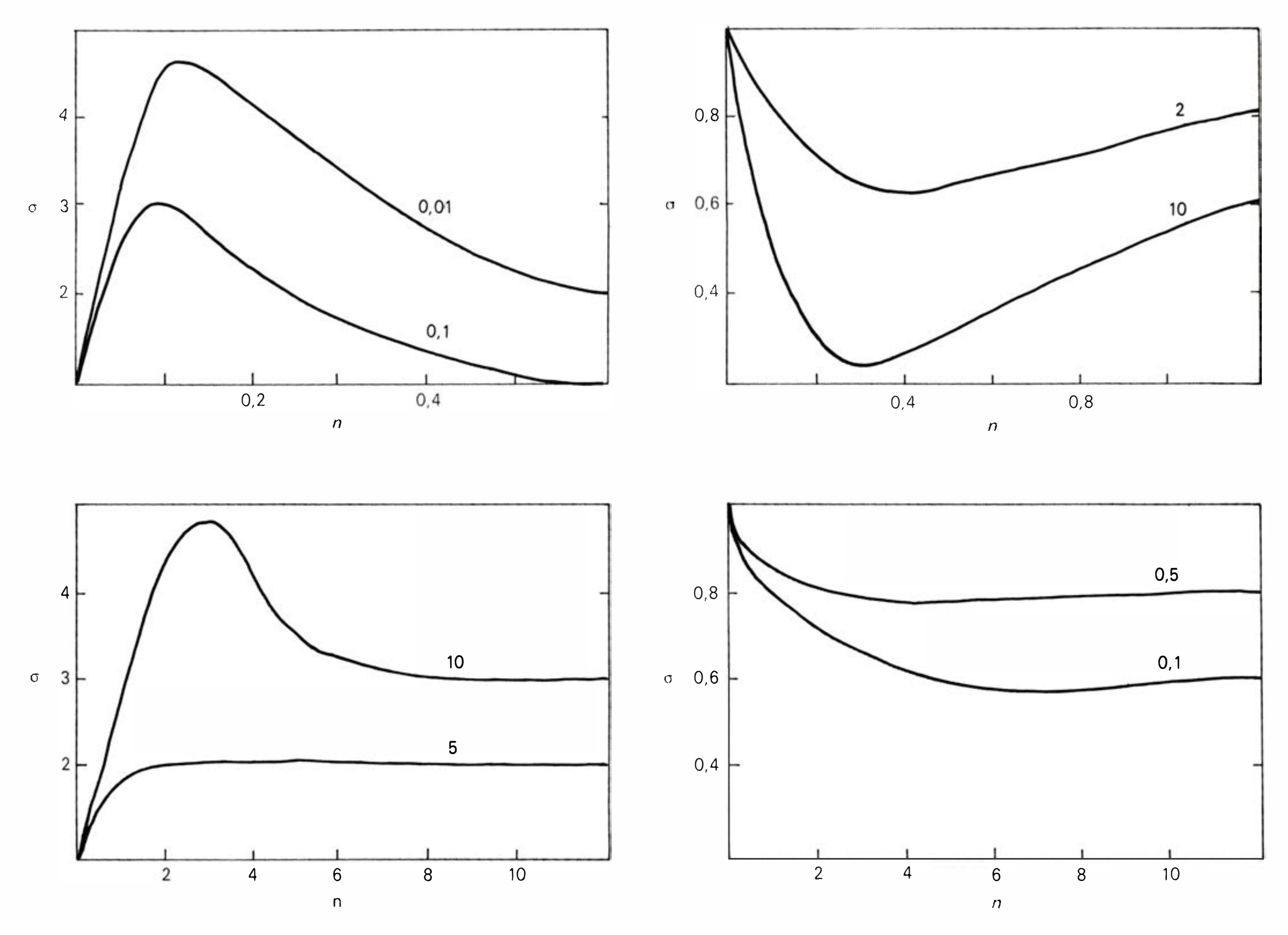
EL ENZIMA ES UN DIMERO que se comporta conforme al modelo alostérico generalizado. En ausencia de ligandos las únicas formas del enzima son las que aparecen en la primera fila de la tabla. Se trata de cuatro confórmeros en equilibrio a través de dos reacciones de transición alostérica. En la primera se equilibran la forma en la primera posición (*dos círculos como reaccionante*) y la forma en segunda o tercera posición indistintamente (*un círculo y un cuadrado*). La correspondiente constante de equilibrio es *L*. En la segunda transición alostérica se equilibran la forma en segunda o tercera posición (*un círculo y un cuadrado*) y la forma en última posición (*dos cuadrados*). La correspondiente constante de equilibrio vale ωL. En presencia del sustrato existe una mayor variedad de formas en equilibrio que difieren entre sí en el grado de ocupación de sus centros de unión.



INFLUENCIA DE LOS VALORES DE LAS CONSTANTES ALOSTERICAS sobre el grado de cooperatividad en la unión con su único sustrato para un enzima que es un dímero generalizado. Para poder apreciar la influencia de los valores de las dos constantes alostéricas se introduce el valor de la relación entre la segunda y la primera, ω, y se hacen las mismas sustituciones que en el caso del dímero concertado. En las gráficas se observa la dependencia de ρ, relación de afinidades a alta y baja concentración del sustrato del valor de la primera constante alostérica para distintos valores de la relación entre las dos constantes alostéricas. En negro se aprecia cómo para valores de ω superiores a la unidad, la cooperatividad en la unión es positiva y tiene un valor máximo tanto más alto cuanto mayor es ω. En la azul puede notarse que para valores de ω inferiores a la unidad la cooperatividad es negativa y tanto más pronunciada cuanto menor sea el valor de ω. (Para cada línea se indica el valor de ω que le corresponde.) En todas, la cooperatividad desaparece cuando *L* es muy grande o muy pequeña comparada con *c*. Para todas, *c* vale 100. Compárese esta ilustración con la figura de la página 115 y las de la página 118.

tivamente sustituye un quimismo por un mecanismo, al hacer desaparecer del modelo alostérico las reacciones acopladas. Pero de momento, también se pueden describir de la forma que hemos indicado más arriba, sin que sea posible preferir una u otra interpretación en todos los casos. De cualquier manera, lo que sí parecen indicar es que los dominios de unión de la molécula proteínica poseen un alto grado de autonomía, como adelantábamos antes.

Los enzimas alostéricos desempeñan un papel extraordinario en la regulación biológica, pero no son las únicas proteínas que intervienen en cometidos de esta índole. En el interior de las células existen otras que, al igual que los enzimas alostéricos, se pueden presentar en dos conformaciones cuyo equilibrio se desplaza como consecuencia de la unión con un ligando específico. Entre ellas están los receptores de las hormonas esteroideas y los represores que se enlazan con el metabolito inductor o represor según los casos. Unas y otras proteínas cambian su especificidad como consecuencia de la unión con sus efectores, modificando sus interacciones con el material genético en regiones específicas del ADN y, de este modo, participan en el control de la síntesis de enzimas específicos. También sobre la membrana celular existen proteínas que constituyen una pieza clave en la comunicación de la célula con el medio y de la célula con otras células del organismo. Son los receptores para sustancias como glutatión en la hidra que provocan en este organismo respuestas en su comportamiento y como los receptores para las hormonas polipeptídicas en los animales superiores que provocan respuestas metabólicas en los tejidos sensibles que los poseen. Precisamente son estos elementos de control los que bien pudieran llevar nuestra compren-



INFLUENCIA DE LOS VALORES de las constantes alostéricas sobre el grado de cooperatividad en la catálisis para un enzima que es un dímero generalizado. Para poder apreciar la influencia de los valores de las dos constantes alostéricas se introduce la relación entre los valores de la segunda y de la primera de ellas, ω, y se hacen las mismas sustituciones que en el caso del dímero concertado. En las gráficas se observa la dependencia de σ (relación entre las eficacias catalíticas a alta y baja concentración de sustrato), del valor de la primera constante alostérica. Como se puede apreciar, la cooperatividad en la catálisis es negativa si la relación entre las constantes alostéricas y la relación entre las eficacias catalíticas de los confórmeros del protómero γ, son a la vez superiores o inferiores a la unidad. Por el contrario es positiva, si una relación es superior y la otra inferior a la unidad. En todas las figuras *c* vale 100. En las de arriba, ω vale 100 y en las de abajo, 0,01. (Sobre cada gráfica se indica el valor de γ.)

sión de los efectos alostéricos literalmente a una nueva dimensión.

Numerosos receptores hormonales específicos situados en la membrana celular ejercen sus características acciones a través de adenosín monofosfato cíclico (AMPc). En la propia membrana existen moléculas del enzima ciclasa de adenilato que cataliza la reacción en que se genera AMPc a partir de ATP. La unión de la hormona con su receptor específico determina la activación de la ciclasa que a su vez provoca un aumento en el nivel intracelular de AMPc. Por su parte este aumento en la concentración de AMPc sirve de estímulo para la activación de enzimas capaces de fosforilar proteínas y con ello poner en marcha variadas respuestas de amplitud progresivamente creciente. Según Pedro Cuatrecasas la activación de la ciclasa resulta de su agregación con las moléculas de receptor ocupadas por la hormona que así pasarían a formar parte de la estructura del mismo enzima. Es probablemente significativo que mediante una técnica que permite obtener celulas híbridas se haya probado que, si una aporta la ciclasa y otra el receptor hormonal, se puede lograr que la célula resultante responda a la hormona con la activación de la ciclasa.

Hay otros receptores de membrana como los específicos para la insulina en el músculo y los que en determinadas neuronas responden a neurotransmisores como acetilcolina que una vez ocupados estimulan la acción de sistemas de transporte específicos situados en la misma membrana que permiten el paso de glucosa o iones. También en este caso es probable que receptor y portador, aunque son entidades distintas, puedan pasar a formar parte de un mismo agregado una vez el receptor se halle ocupado.

Es digno de mención que para algunos receptores de sustancias activas tanto exógenas como endógenas y tanto en invertebrados como en vertebrados hay pruebas de que la unión con la sustancia activa presenta cooperatividad negativa. En estos casos es tentador suponer que también el propósito del efecto homotrópico sea impedir en buena medida que se altere el grado de ocupación de los receptores cuando se produzca una disminución en el nivel de la sustancia activa, sobre todo si se tiene en cuenta que sus niveles son generalmente muy bajos. De hecho para muchos receptores no sólo se conoce la existencia de efectos homotrópicos sino también parece probado que determinadas sustancias específicas modulan la respuesta que resulta de la ocupación del receptor. Sin duda, si se llega a probar en todos sus detalles la existencia de oligómeros en la membrana celular, como parece próximo a lograrse con el receptor para acetilcolina, el modelo alostérico entrará en un nuevo ámbito bidimensional.

Juegos matemáticos

El juego de damas, más interesante de lo que muchos podrían pensar

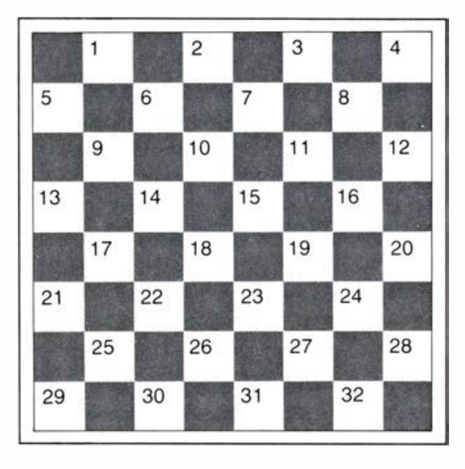
Martín Gardner

"El juego de damas que conocemos está especialmente calculado para fijar la atención sin fatigarla. Hay en las damas seriedad y compostura, que insensiblemente van apaciguando el espíritu."

JAMES BOSWELL,
The Life of Samuel Johnson

La cita ha sido tomada de una sección dedicada al año 1756, en la que Boswell comenta el prefacio que Johnson puso a una obra de William Payne: *Introduction to the Game of Draughts*, publicada en Londres ese mismo año. Este libro, escrito por un profesor de matemática, fue el primero escrito en inglés acerca del juego de damas. Johnson apenas si jugó a él después de abandonar la universidad, lo que Boswell lamenta, porque opina que jugando a las damas su amigo hubiera podido tener "sedante e inocente alivio" en sus periódicos trances de depresión.

Nada se sabe sobre sus orígenes, aunque casi todos los especialistas en historia de los juegos coinciden hoy que se originó en el sur de Francia, en algún momento del siglo XII. En Inglaterra y Estados Unidos es, sin duda, el más conocido de los juegos de tablero; basta pensar en la cantidad de niños que aprenden a jugar, quienes ya no olvidarán las reglas. Y aunque la bibliografía sobre damas es, al menos en volumen, muy inferior a la ajedrecística, no olvidemos el gran número de adultos que llegan a ser jugadores de primera, ni el interés que suscitan en el público los torneos del campeonato mundial de damas. ¿Cuánta gente sabría nombrar al menos un experto en damas, o decir quién es el actual campeón del mundo? Tal gloria le cabe al doctor Marion F. Tinsberg, especialista en topología, del Departamento de Matemática de la Florida A. and M. University, y, probablemente, el mejor jugador de damas que haya existido jamás.



Numeración de las casillas del tablero de damas

En el llamado mundo occidental, las reglas de ajedrez ya están normalizadas, pero no sucede lo mismo con las damas. (En España, Inglaterra y Estados Unidos se juega sobre un tablero de 64 escaques; cada jugador dispone de 12 piezas, llamadas peones, situados sobre los cuadros negros de las tres primeras filas contando a partir de él; las piezas solamente pueden moverse sobre los cuadros negros. Nota del traductor.) Fuera del mundo de habla inglesa existen docenas de variantes regionales. La versión más difundida en Europa y la Unión Soviética, llamada "damas polacas", excepto en Polonia, donde se llaman "damas francesas", se juega sobre un tablero de 10 por 10, y cada bando dispone inicialmente de 20 peones. Esta es la versión normal del juego en Francia. En el Canadá francófono el tablero es todavía mayor: de 12 por 12, con 30 piezas por bando. Las reglas de las damas difieren ampliamente de unos a otros lugares del mundo. Y es curioso que en todos los países europeos, excepto Gran Bretaña, las piezas suelen llamarse "damas", mientras que en Estados Unidos y demás países anglófonos se llaman "hombres".

Del hecho de ser las damas juego más sencillo que el ajedrez se desprenden varias consecuencias. Entre ellas, que resulta mucho mas improbable que un jugador de nivel "gran maestro" pueda perder una partida frente a un oponente inferior debido a un error de lo que sucede con su equivalente ajedrecístico. Para los entusiastas, este es uno de los grandes atractivos del juego. Y les encanta una cita de Edgar Allan Poe, cuando éste compara ajedrez y damas al comienzo de *The Murders in the Rue Morgue:*

"Por consiguiente, haré uso de esta oportunidad para declarar que las potencias superiores del intelecto reflexivo son más decidida y fecundamente empleadas en el nada ostentoso juego de damas que en el ajedrez, tan refinadamente frívolo. En este juego, donde las piezas tienen movimientos *extravagantes* y distintos, y valores diversos y variables, se toma como profundo lo que solamente es complicado (error nada infrecuente). En él, se solicita poderosamente la *atención*. Un solo instante que decaiga, y ya se ha cometido un descuido, que resulta en daño o en derrota. Al ser las jugadas posibles, no sólo numerosas, sino intrincadas, tales ocasiones se multiplican, y de cada diez casos, en nueve no es el jugador más perspicaz quien vence, sino el capaz de mayor concentración. Por el contrario, en las damas, donde los movimientos son *únicos* y no tienen sino pequeña variación, las probabilidades de fallos inadvertidos son mucho menores, y estando la mera atención comparativamente poco solicitada, las ventajas que se consigan por cada bando serán obtenidas gracias a mayor *perspicacia*."

Tinsley lo ha expresado de este modo: "jugar al ajedrez es como contemplar un océano sin fin; jugar a las damas es como mirar un pozo sin fondo."

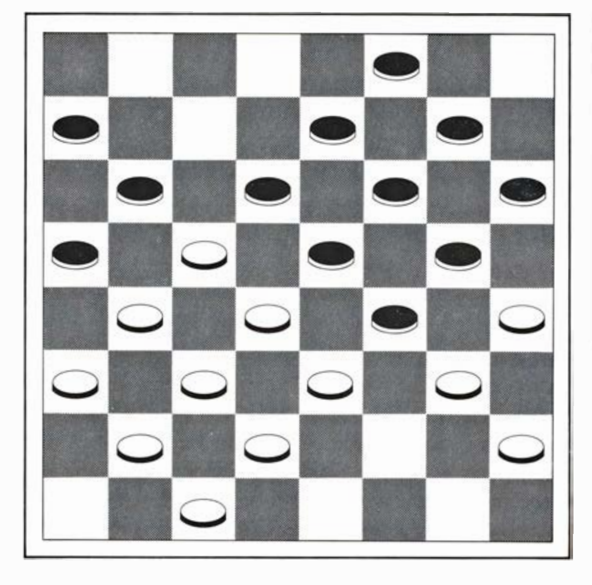
Otra de las consecuencias de la mayor sencillez de las damas es que hacia 1900 las posibles aperturas del juego habían sido todas tan exhaustivamente estudiadas que no había torneo que no concluyera en tablas. Para hacer la competición un poco más apasionante, se introdujo en Gran Bretaña la práctica de anotar en tarjetas combinaciones de un par de movimientos: el de apertura de las negras, y la respuesta de las blancas, ambos obligatorios para los jugadores.

Dado que cada bando puede optar entre siete posibles aperturas, hay 49 pares posibles. Dos de ellos (9-14, 21-17 y 10-14, 21-17) fueron desechados porque sacrifican al contrario una pieza blanca. Más tarde se vio que otros dos pares más (11-16, 23-19 y 12-16, 23-19) le dan a las negras tan fuerte ventaja sobre las blancas, que tuvieron que ser descartados también. Quedaron así, pues, 45 tarjetas.

La notación habitual de las partidas de damas se basa en la numeración de casillas que se ve en la ilustración de la página anterior. Para mayor claridad, los diagramas posicionales acostumbran a dibujarse invirtiendo los colores de los cuadros, ocupando las piezas casillas blancas en lugar de negras; pero el verdadero juego se desarrolla siempre sobre los cuadros negros, con el "ángulo doble" situado del lado inferior derecho de cada jugador. Es costumbre llamar a los contrincantes Blanco y Negro, aun en el caso de que las piezas sean blancas y rojas. Modernamente los torneos se juegan sobre tableros de colores verde y beige claro; los tableros negros y rojos —alguno hay— son considerados atrocidades de almacén de juguetería. El primero en mover es siempre el bando negro; las partidas se anotan con las negras ocupando las casillas de numeración más baja en el momento de comenzar. Si tiene usted la intención de trabajar en alguno de los pasatiempos de esta sección, será buena idea rotular los cuadros negros de su tablero como aquí se muestra.

Es lástima, pero con el paso de los decenios los especialistas llegaron a familiarizarse tanto con todas las variantes que emanan de las aperturas obligatorias, que pronto encontraron sistemas de juego "seguro", con lo que se retornó a las series de partidas terminadas en tablas. La restricción inglesa "de dos jugadas" se sustituyó en los Estados Unidos por tres jugadas obligatorias, ya a mediados de los años 30, práctica que actualmente es adoptada en casi todos los torneos norteamericanos y británicos. Hay ahora 142 tarjetas; cada una contiene una terna distinta de jugadas obligatorias. Puesto que muchas de estas ternas conceden ventaja a uno de los bandos, por lo común al segundo jugador, con cada salida obligatoria se juegan dos partidas, para que cada jugador tenga así derecho a mover en primer lugar.

Si no fuera por estas condiciones de apertura obligatoria, los especialistas actuales no conseguirían sino hacer tablas. Incluso con la obligatoriedad de las tres primeras jugadas, alrededor del 80 por



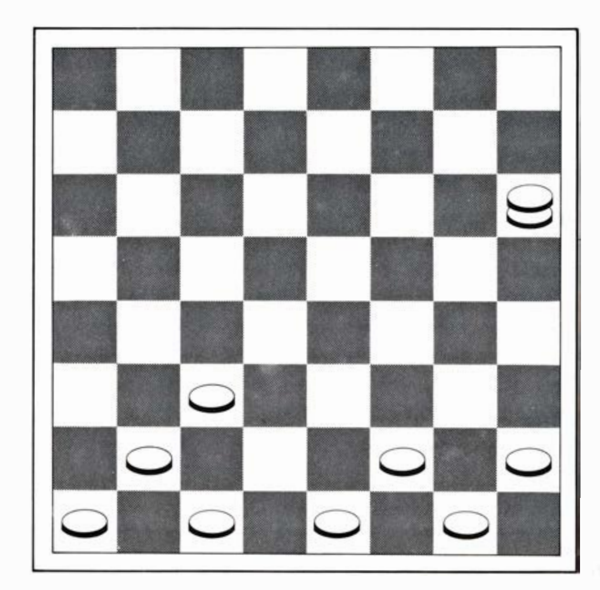
	BLANCAS	NEGRAS
1.	9-13	24-20
2.	12-16	21-17
3.	10-15	23-18
4.	15-19	18-14
5.	8-12	25-21
6.	4-8	29-25
7.	6-10	27-23
8.	10-15	23-18
9.	2-6	31-27
10.	6-9	27-24
11.	1-6	32-27
12.	6-10	27-23

La partida de damas más breve sin capturas

ciento de las actuales partidas de torneo acaban en tablas. Por lo general, cuando uno de estos grandes jugadores consigue vencer es porque su contrincante hizo una patochada, o porque el vencedor consiguió mantener en secreto (algunas veces durante años) un "guiso" o "receta" descubierta por él. Lo mismo que en ajedrez, los "guisos" son modificaciones y perfeccionamientos de "técnicas de manual" que toman por sorpresa al contrario. Tradicionalmente, a los jugadores se les concedían cinco minutos para pensar cada jugada, y un minuto para aceptar capturas que solamente fuesen realizables de una manera. En estos últimos años tal práctica ha sido reemplazada por relojes de ajedrez, estando los jugadores obligados a 30 movimientos en un máximo de una hora. Cuando alguien se presenta con un guiso nuevo, lo que ocurre, sencillamente, es que su víctima carece de tiempo para analizarlo.

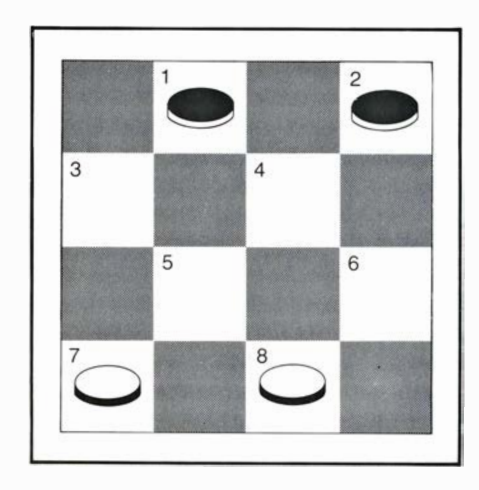
En 1967, Walter Hellman, ya fallecido, obrero siderúrgico de Gary, Indiana, que a la sazón era campeón del mundo, defendió su título frente al campeón de los Estados Unidos, Eugene Frazier. El torneo fue a 36 partidas, de las cuales 31 acabaron en tablas, y cinco fueron victorias de Hellman. La última de estas victorias se debió a un guiso. "Ya había utilizado esta receta en una ocasión anterior", declaró Hellman a un periodista, "pero no ha sido publicada jamás. Frazier tenía tan sólo una jugada que pudiera bloquear mi ataque, y cinco minutos no dan para mucho".

Una tercera consecuencia de la sencillez de las damas es que los mejores programas computarizados de este juego presentan a los jugadores de nivel medio dificultades mucho más formidables que



	BLANCAS	NEGRAS
1.	11-16	21-17
2.	10-14	17x10
3.	6x15	23-18
4.	2-6	18x2 (K)
5.	9-14	2x18
6.	3-7	24-20
7.	1-6	20x2 (K)
8.	12-16	2x9
9.	5x23	26x3 (K)
10.	4-8	3x12

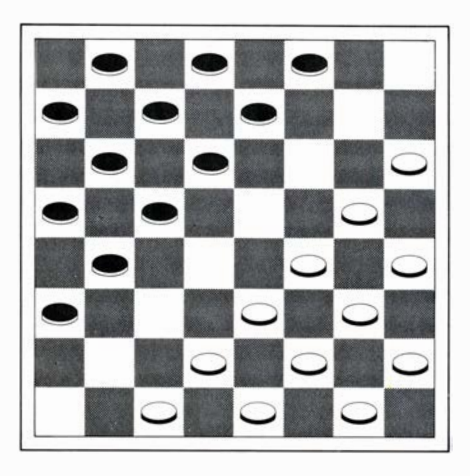
La partida más breve hoy conocida



Un rompecabezas de mini-damas

los correspondientes programas de ajedrez. Hasta hace unos pocos años, el programa más temible era resultado de los trabajos de Arthur L. Samuel. Se trata de un programa capaz de "aprender", cuya eficacia mejora conforme juega. Tras retirarse de su cargo de director de investigación en IBM, Samuel ha continuado perfeccionando su programa en el Laboratorio de Inteligencia Artificial de la Universidad de Stanford. En estos últimos años ha sido proyectado y puesto a punto un potente programa –no del tipo de aprendizaje–, por Eric C. Jensen y Tom R. Truscott, ambos estudiantes de segundo ciclo en la Duke University, quienes trabajan con un especialista en inteligencia artificial, Alan W. Biermann. Se cree que este programa es superior al de Samuel, aunque esto todavía no ha quedado zanjado en un torneo de manera concluyente.

Los jugadores de damas están encuadrados en tres niveles, que en los Estados Unidos se llaman minor, major y master (maestro). Los partidarios del programa Duke creen que inicialmente



Damas en diagonal

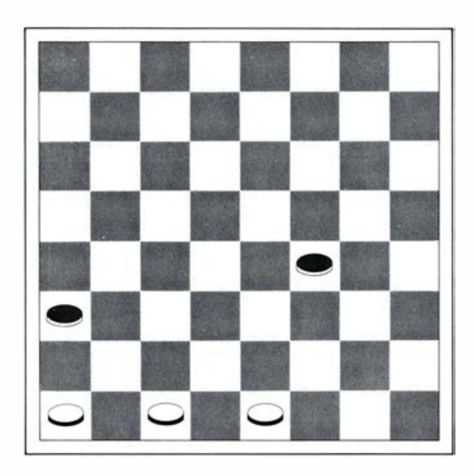
juega al nivel de maestro. No obstante, tras jugar algún tiempo contra el programa, los grandes maestros consiguen apreciar sus puntos débiles y beneficiarse de ellos. Su mayor debilidad es que juega sin plan preconcebido; ni siquiera sigue las indicaciones de los manuales para las jugadas de apertura, y por lo general distribuye las piezas propias por el tablero en configuraciones que los maestros consideran estúpidas. Su fuerza reside en la increíble velocidad con que puede analizar todas las jugadas posibles hasta niveles mucho más profundos que sus contrincantes humanos; y ello sin cometer error ninguno. Es posible que se necesiten todavía varios decenios para conseguir programas de ajedrez de nivel de gran maestro, pero en opinión de Biermann, el programa Duke ya "está llamando a la puerta" del campeonato del mundo. Jensen y Truscott están convencidos de que su programa podría colocarse hacia el décimo lugar en la clasificación mundial.

Los jugadores de nivel "gran maestro" tienen en cambio conceptuación mucho menos entusiasta sobre la calidad de los programas de ordenador. Todos ellos coinciden con W. Burke Grandjean, secretario de la American Checker Federation, quien considera que el optimismo del grupo de la Duke es "ingenuo hasta el ridículo". Respaldado por la Federación, Tinsley mantiene 5000 dólares de apuesta a que podrá vencer en un torneo de 20 partidas a cualquier programa que pueda idearse dentro de los próximos cinco años. Biermann piensa que en la actualidad Tinsley acabaría por vencer al programa Duke, pero no sin antes ceder unas cuantas partidas. Grandjean afirma poder encontrar fácilmente 20 jugadores dispuestos a enviar 100 dólares cada uno, que apostarían a que Tinsley no perdería ni siquiera una partida, y por su parte, él mismo está persuadido de que Tinsley vencería en todas y cada una de las 20 partidas. Quizá podamos saberlo pronto. (Los lectores interesados en ponerse en contacto con la American Checker Federation, y en recibir su *Bulletin*, mensual, pueden escribir a Grandjean, al 3475 Belmont Avenue, Baton Rouge, La., 70808.) En la actualidad, Fidelity Electronics tiene en el mercado una máquina de estado sólido y bajo precio, el Checker Challenger 2, capaz de jugar a dos niveles de dificultad, y también, el Checker Challenger 4, capaz de hacerlo a cinco niveles; el nivel más alto de estas máquinas está considerado como inferior al de los programas de Samuel y de la Duke University.

En ajedrez es fácil demostrar que la partida más breve posible es el "mate del tonto", donde el segundo jugador alcanza el mate en su segunda jugada. Aunque parezca curioso, se desconoce cuál sea la partida de damas más breve posible. Hasta hace dos años se pensaba que sería la conducente al juego bloqueado que se ve en la ilustración superior de la página precedente, que para alcanzar esta posición final consume 24 jugadas. Hay muchas secuencias de 24 jugadas capaces de llevar a esta situación, pero se cree que la posición final es en realidad única. Con el desarrollo de juego expuesto, cada jugada de las blancas es la simétricamente opuesta de la precedente de las negras con relación al centro del tablero. Desconozco quién dio a la partida forma simétrica por vez primera. La versión que aquí expongo, preparada por Rudolf Ondrejka, de Linwood, N.J., comienza por la apertura Edinburgh en dos movimientos. Dado que 9-13 (una de las jugadas favoritas de todos los bisoños) está reputada como una de las peores aperturas posibles de las negras, el juego suele comenzar por 10-15, 23-18, apertura conocida por Kelso Cross.

Sam Loyd, en su *Cyclopaedia of Puzzles* (1914), página 374, sirviéndose de una extravagante notación que incorrectamente presume que el tablero ha sido girado 90 grados, recoge una secuencia asimétrica de 24 jugadas que termina en la configuración antes mencionada. Loyd declara lisa y llanamente que se trata de "la mínima partida posible". En realidad, la partida bloqueada de 24 jugadas es la más breve de las partidas donde no haya capturas, y así puede demostrarse. No obstante, en 1978, Alan Malcolm Beckerson, redactor jefe de *English Draughts Journal* descubrió que las blancas podrían vencer en su décima jugada (con un total de 20 movimientos en la partida) ¡capturando todas las piezas negras! Si bien ésta es la partida de damas más breve que hoy se conoce, nadie ha podido demostrar que sea imposible lograr otra más corta. Beckerson ha descubierto varias partidas más, todas de 20 jugadas, donde son capturadas todas las piezas negras, o donde, tras algunas capturas, se llega a un juego bloqueado. La versión que vemos en la ilustración inferior de la página precedente, donde el tablero muestra la posición final, se publicó por primera vez en una revista mensual inglesa, *Games and Puzzles* (marzo de 1978). Sus dos primeras jugadas se conocen como apertura de Newcastle.

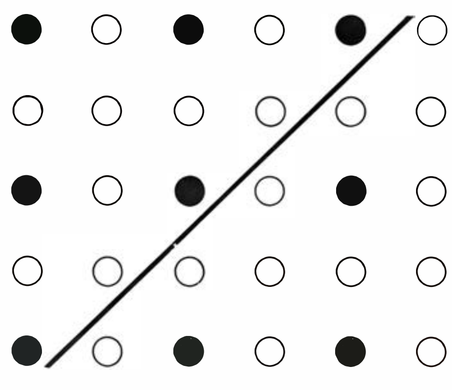
En las damas hay otros muchos pro-



Apuesta engañabobos

blemas "de mínimo" que distan de estar resueltos. ¿Cuál es el mínimo número de jugadas lícitas para que una partida produzca 24 damas? La mejor de las soluciones conocidas, en 90 jugadas, apareció en *Journal of Recreative Mathematics* (vol. 9, n.º 1, pág. 45; 1976). ¿Y cuál es el mínimo número de jugadas para que blancas y negras inviertan sus posiciones iniciales? Si el tablero estuviera ocupado tan sólo por un bando se necesitarían al menos 60 jugadas para ocupar las casillas que inicialmente pertenecerían al bando contrario; por consiguiente, una cota inferior serían 2 por 60, o sea, 120 jugadas. Puede verse una solución formada por 172 jugadas en una obra inglesa de finales del siglo pasado, *The Draughts-Players Guide and Companion*, por Frank Dunne, págs. 94-95. Al terminar, cada bando habrá conseguido coronar seis peones. Parece razonable que las 172 jugadas podrían rebajarse considerablemente.

Vale la pena intentar el problema en tableros más pequeños. En el de 3-por-3 el problema es trivial; pero en el caso 4-por-4 se presenta ya un entretenido rompecabezas. Partiendo de la posición que vemos en la figura superior de la



Una solución "¡Ajá!"

página 122, la tarea consiste en intercambiar ambos bandos con el mínimo número de jugadas lícitas. Como es evidente, las capturas son obligatorias. Al terminar, las cuatro piezas habrán sido coronadas. Daré una solución el mes que viene. De pasada, la partida mínima exige en este tablero cinco jugadas. Si ambos bandos juegan a ganar y aplicar a tal fin la estrategia óptima, el juego termina en tablas.

Lo mismo que en ajedrez, se han propuesto para las damas infinidad de variantes, ya sea modificando el tablero, ya las posiciones iniciales o las reglas de juego. Un libro francés, *Les jeux de dames non orthodoxes et autres jeux à pions* publiçado en 1956 en París por cuenta de sus autores, Joseph Boyer y Vern R. Parton, presenta más de 100 de estas variantes. Algunas se juegan sobre teselaciones triangulares o hexagonales; las hay incluso en tableros tridimensionales. En otras, las piezas se mezclan con las de ajedrez; algunas permiten competir a tres o cuatro jugadores simultáneamente. Como es fácil imaginar, resulta imposible trazar una divisoria nítida entre juegos lo bastante parecidos a las damas como para ser llamados variantes, y lo suficientemente distintos como para ser considerados como juegos enteramente distintos. Así, por ejemplo, las llamadas "damas turcas" apenas guardan semejanza con las ordinarias, salvo en que se juegan sobre un tablero de 8-por-8 con fichas de dos colores. Una sencilla variante de las damas normales consiste en situar inicialmente las piezas como se ve en la ilustración inferior de la página 122. Todas las reglas de damas siguen siendo válidas. Las extravagantes situaciones que se producen son consecuencia de estas posiciones iniciales, y no se encuentran en ninguna partida de damas ortodoxa.

Quizá la más excéntrica variante de las damas, sobre la que nos gustaría saber más, sean las "superdamas", inventadas por Charles Fort. (Fort vivió en el Bronx. Recopilador de toda suerte de anomalías científicas, ejerció marcada influencia sobre la literatura de ciencia-ficción, y es responsable de buena parte del epidémico interés actual por todo lo paranormal.) Según el biógrafo de Fort, Damon Knight, autor de *Charles Fort* (Doubleday, 1970), las superdamas se jugaban "con ejércitos de peones, sobre un tablero de miles de casillas. Fort usaba trozos de cartulina como peones, con una chincheta para asirlos, y como tablero, una pieza de tela a cuadros".

Los dos jugadores empiezan por situar sus piezas en cualquier formación preconvenida que deje espacio entre ambos ejércitos. Si los jugadores solamente pudieran mover una pieza por vez la partida podría durar semanas; por ello, Fort permitía ataques en masa. He aquí como lo expresó él mismo en una carta: "Supongamos que *A* empieza el primero; va moviendo piezas hasta que *B* le mande parar –pongamos por caso, cien jugadas–. Entonces *B* realiza cien jugadas. Puede que después *A* desee volver a efectuar otras cien jugadas, pero entonces *B*, calibrando la situación, puede mandarle parar, por ejemplo, al llegar a treinta. Quizá se produzca entonces un cuerpo a cuerpo, con una jugada por vez, como en las damas corrientes. Pero si en un momento dado uno de los jugadores desea hacer un movimiento masivo, todo lo que necesita es autorización del contrario".

Por lo general una partida duraba toda la noche. En 1930, Fort le escribió a Tiffany Thayer, editora de *Doubt*, la primera revista sobre temas "fortianos": "Las superdamas van a ser un gran éxito. Acabo de conocer a otras cuatro personas que han considerado el juego una completa estupidez".

En Inglaterra y los Estados Unidos la variante más popular es la llamada "giveaway", "juego a traición", cuya única diferencia con el ordinario consiste en que, para ganar la partida, es necesario no tomar todas las piezas del contrario, sino desembarazarse de todas las piezas propias. En el libro de Dunne ya citado, páginas 91-92, hay una fantástica "apuesta engañabobos", presumiblemente preparada por engañabobos profesionales. El bando blanco comienza con los 12 peones situados en la posición inicial ordinaria; las negras disponen únicamente de una dama situada en la casilla 7. El negro gana si consigue perder su dama. El blanco gana si consigue obligar al negro a tomarle sus doce peones. Dunne muestra que las blancas pueden vencer siempre; y da otras tres apuestas parecidas, donde el negro comienza con un solo peón en la casilla 1, la casilla 4 o la casilla 5.

Entre los cientos de apuestas que los engañabobos han amañado, una de las mejores parte de la posición que vemos en la parte superior de esta misma página. (Estoy en deuda con Mel Stover, quien me lo ha dado a conocer). Les toca mover a las negras. Las blancas apuestan a que las negras no serán capaces de coronar la pieza que muevan en primer lugar. Como es obvio, las negras no deben mover su peón de la casilla 21, pues lo perderían inmediatamente; así pues, la cuestión es si las negras pueden

mover la pieza del cuadro 19 y hacerla avanzar hacia su línea de coronación. Cuanto más se estudia el problema, tanto más evidente resulta que las negras pueden ganar la apuesta sin ninguna dificultad. Y no obstante, son las blancas quienes ganan. Puede resultar divertido apostar con los amigos. Si el lector no consigue descubrir la estrategia de las blancas (que no tiene trampa ni cartón) podrá ver la solución el mes que viene.

Un último problema. Está generalizada la opinión de que dos damas pueden siempre vencer a una sola dama; empero no es así en todos los casos. Vea el lector si es capaz de situar dos blancas y una negra de forma que siendo el turno de las blancas, el bando negro pueda forzar tablas.

De los problemas del mes pasado, el primero fue resuelto ya dentro de aquella sección; las respuestas a los restantes son como sigue:

2. El misterio del póker se resuelve inmediatamente notando que en una mano que contenga cuatro naipes de igual valor ("póker") hay una quinta carta. Por tanto, para cada póker hay 48 distintas quintas cartas posibles; por ello hay 48 × 13, o sea, 624 maneras de conseguir una mano con cuatro naipes del mismo valor, frente a tan sólo 40 posibles escaleras de color. Esta es la razón de que sea mucho menos probable que en el reparto se reciba una escalera de color que un póker, y en consecuencia, la escalera de color tiene más valor que el póker. El problema fue comunicado por M.H. Greenblatt al *Journal of Recreational Mathematics* (Vol. 5, n.º 1, pág. 39; enero de 1972).

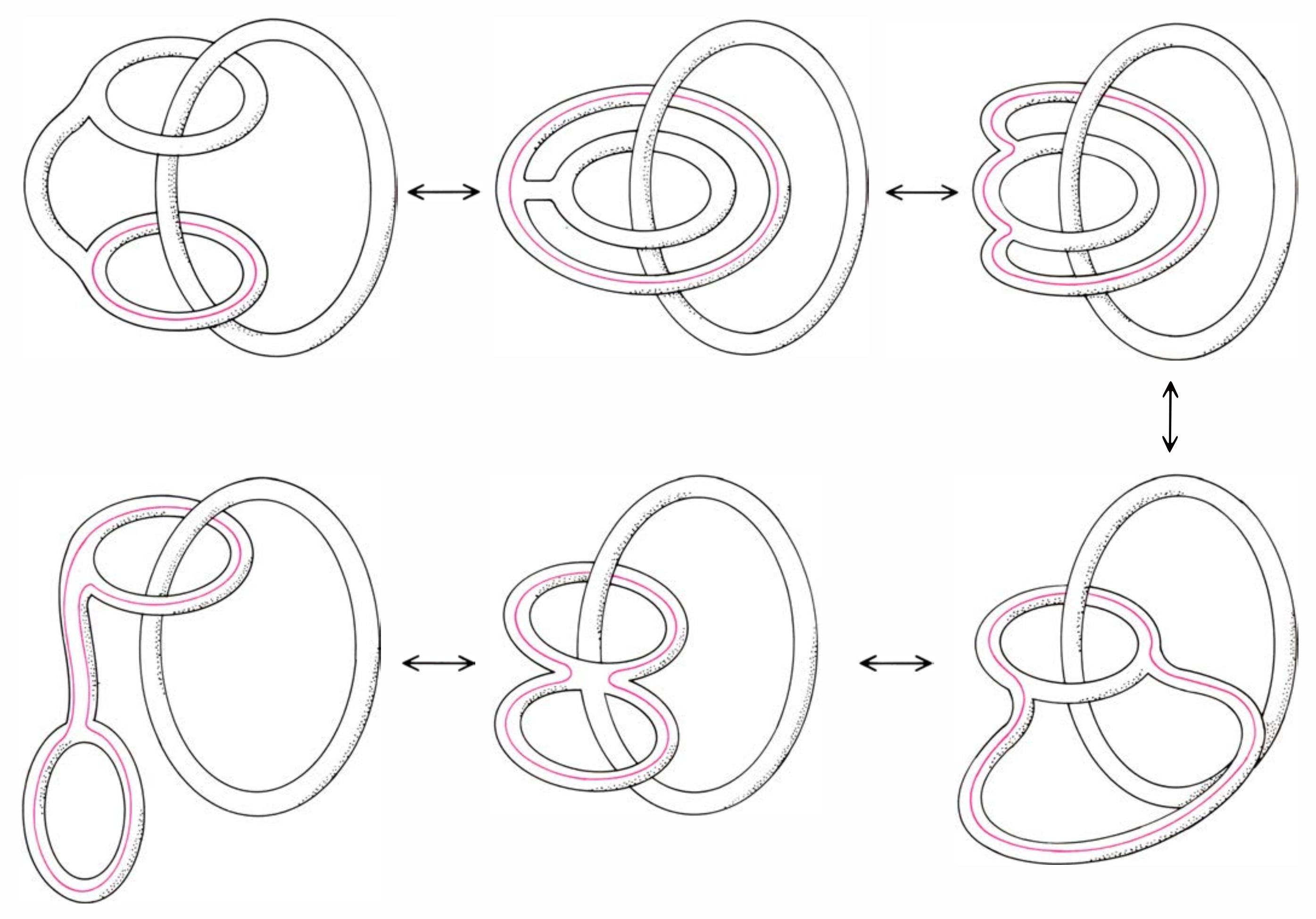
3. He aquí cómo Raymond Smullyan, en su nuevo libro *The Chess Mysteries of Sherlock Holmes* (Knopf, 1979) demuestra que con la posición de juego mostrada el mes pasado, las piezas verdes fueron las primeras en jugar.

"Las rojas están en jaque; luego el último en jugar fue el bando verde. Queda por determinar quién movió primero, lo que puede conseguirse calculando si el número de jugadas realizadas es par o impar.

"La torre situada en *b* 1 ha efectuado un número impar de movimientos; las otras tres torres han hecho un número par de jugadas (posiblemente, cero). Los caballos rojos han hecho en conjunto un número non de saltos, pues al saltar el caballo cambia el color del cuadro, y ambos están en cuadros del mismo color, mientras que los caballos verdes han realizado colectivamente un número par. Uno de los reyes ha intervenido en un número par de jugadas (posiblemente, cero) y el otro, en un número impar. Los peones no han sido movidos, y los alfiles, tampoco. Las damas han sido tomadas antes de que pudieran mover. Así pues, el total general es impar. Por tanto, las verdes movieron en primer lugar. Las verdes son blancas, y las rojas, negras."

4. Aunque pueda parecer sorprendente, tanto las clases más ricas como las más pobres de Petrolandia prefieren que se promedie por pares empezando por las clases más adineradas. Los más ricos preferirán promediar sus bienes con la clase inmediata antes de que esta última pierda parte de su riqueza al promediar. Los de la clase más pobre preferirán también promediar con los de la clase inmediatamente superior después de que éstos, al promediar, se enriquezcan algo.

Un ejemplo numérico lo aclarará del todo. Supongamos que la riqueza de las



Solución del problema de los toros entrelazados

cinco clases guarde la proporción 1:3:4:7:13. Promediando desde la clase más pobre hacia la más rica resultan las proporciones 2:3:5:9:9, mientras que al hacerlo en sentido contrario las proporciones cambian a 3:3:5:7:10.

5. Dividamos los 500 kilómetros de línea férrea en 10 tramos de 50 kilómetros cada uno. Si alguno de éstos es recorrido en una hora el problema está resuelto; por consiguiente, debemos suponer que cada uno de ellos es recorrido ya sea en más, ya sea en menos tiempo. Se deduce que en algún lugar de la vía existirán un par al menos de tramos adyacentes donde uno (que llamaremos *A*) se recorre en menos de una hora, y el otro (*B*) lo es en más.

Imaginemos ahora una enorme varilla métrica de 50 kilómetros de largo, colocada sobre el tramo *A*. Vayamos mentalmente corriendo la varilla en la dirección del tramo *B*, hasta dejarla superpuesta a *B*. Conforme la varilla va deslizándose, el tiempo requerido por el tren para recorrer estos 50 kilómetros va variando continuamente, desde menos de una hora (tramo *A*) hasta más de una hora (tramo *B*). Por consiguiente, debe existir al menos una posición donde la varilla cubra un tramo de 50 kilómetros recorrido por el tren en una hora exactamente.

6. La feliz idea que resuelve el problema de "¡Ajá!" consiste en pintar nueve puntos de color, como se muestra en la ilustración inferior de la página 124. Es obvio que, no importa como se salte, una moneda colocada en un punto de color solamente podrá llegar a otro punto de color.

Hay seis puntos de color situados por encima de la línea y solamente tres debajo de ella. Por el principio del casillero, quedarán por encima de la raya tres monedas sin sitio adónde ir bajo ella. La tarea de trasladar todas las monedas a puntos situados bajo la raya no podrá llevarse a cabo a menos que tres de ellas, situadas sobre puntos de color, sean retiradas de la formación triangular superior. Quitando tres cualesquiera de éstas, es tarea fácil trasladar las doce restantes.

7. La ilustración inferior antedicha muestra cómo una deformación continua del toro de dos agujeros permite soltar uno de los eslabones que lo encadenan al toro de un solo agujero. El razonamiento de imposibilidad que se dio el mes pasado falla, porque al pintar un anillo alrededor de un agujero (como muestra la línea de color) el anillo se deforma de manera tal que aún liberado un eslabón, el anillo de color sigue encadenado al toro de un solo agujero.

Taller y laboratorio

Espectrofotómetro casero que examina el espectro en un treintavo de segundo

Jearl Walker

En esta sección se han presentado diversos diseños de espectroscopios, espectrofotómetros y demás instrumentos válidos para descomponer la luz en sus colores (o longitudes de onda) componentes, permitiendo así su análisis. Los instrumentos eran básicamente de dos tipos; uno para estudiar las propiedades de absorción luminosa de un líquido, sólido o gas transparente, y otro para analizar una fuente de luz, una lámpara o una estrella por ejemplo. Dean Morelli, de Rye, Nueva York, me acaba de enviar un nuevo diseño para un tipo especial de espectofotómetro. Su principal ventaja reside en la velocidad con que barre el espectro visible. Con él se puede obtener un conjunto completo de medidas en menos de un treintavo de segundo.

Para explicar mejor en qué consiste el aparato de Morelli, echaremos un vistazo a los dos tipos fundamentales de espectroscopio y espectrofotómetro. En el primero de ellos, se hace pasar la luz blanca, procedente de una lámpara, a través de una celdilla que contiene el líquido, el sólido o el gas que queremos investigar. La muestra absorbe la luz en unas determinadas longitudes de onda; y así, esa luz disminuye. Se descompone luego (separándola en longitudes de onda) mediante un prisma o una rejilla de difracción antes de incidir sobre un detector. El detector nos permite determinar qué colores de la luz inicial han sido extraídos por la muestra. El observador sabrá identificar los componentes de una muestra y distinguir una muestra de otras que, aunque a simple vista sean iguales, difieran en su composición, si se ayuda de algún manual de espectroscopía molecular y atómica.

El segundo tipo básico de esos instrumentos suele ser de uso familiar en trabajos de astronomía. El espectroscopio se incorpora al telescopio para que el observador defina los colores presentes en la luz procedente de un objeto celeste. Muchos aficionados han analizado, con este procedimiento, la luz solar. El sol emite radiaciones luminosas a lo largo de un rango de longitudes de onda, con la máxima intensidad cerca del centro del espectro visible. Cuando dicha luz, básicamente blanca, atraviesa la atmósfera, los elementos allí presentes absorben la luz de unas longitudes de onda características. Las longitudes de onda absorbidas aparecen como miles de líneas oscuras (las líneas de Fraunhofer), superpuestas a los colores del espectro visible. Con una frecuencia adecuada, el observador logrará identificar las líneas con los elementos presentes en la atmósfera solar.

Un espectroscopio cuenta con el ojo humano para examinar el espectro. Si un aparato sensible a la luz, piénsese en un fotomultiplicador, sustituye al ojo, el instrumento será un espectofotómetro. El experimentador todavía debe ir tomando lecturas de la intensidad de la luz, punto por punto, a lo largo del espectro. Un sistema mejor aún consiste en incorporar al fotomultiplicador un registrador, de suerte que, cuando el fotomultiplicador recorra el espectro, su respuesta quede automáticamente registrada en un papel en forma de curva de intensidad para cada longitud de onda.

Registrar la intensidad de la luz punto por punto, a lo largo del espectro, puede requerir bastante tiempo. Las longitudes de onda se miden en angstrom; el espectro visible tiene 3000 angstrom de ancho. Por tanto, para hacer una medida cada 100 angstrom, será necesario efectuar 31 mediciones. Ello nos llevaría muchos minutos si quisiéramos cubrir el espectro completo. El espectofotómetro de registro rápido de Morelli puede hacer ese trabajo en una fracción de segundo dando, por tanto, al experimentador la oportunidad de investigar los fenómenos de manera mucho más rápida de como lo haría con instrumentos más lentos.

Como ejemplo de capacidad de rapidez de observación, Morelli me contó que él mismo había examinado las oscilaciones de color del oscilador químico que describimos en nuestra colaboración del mes de septiembre de 1978. Aunque el período de oscilación depende de las concentraciones de las sustancias químicas presentes, no suele llegar al minuto. Con un espectofotómetro convencional uno no se puede dar por satisfecho con seguir las oscilaciones en una sola longitud de onda, ya que el aparato nunca le permitirá mantener la velocidad con que se producen las reacciones.

El diseño del instrumento de Morelli puede modificarse en función de las necesidades del experimentador. Pero en líneas generales consta de una lámpara, una red de difracción curvada, un tubo fotomultiplicador y un osciloscopio. Se hace pasar la luz de la lámpara a través de una célula que contiene la muestra a investigar. La red de difracción dispersa la luz en las diferentes ondas que la componen. La red puede desplazarse horizontalmente con el fin de que el tubo fotomultiplicador rastree todo el espectro visible; la intensidad de la luz viene indicada en el osciloscopio. En el eje vertical del osciloscopio se representan las intensidades y, en el horizontal, las longitudes de onda. Por tanto, se ve un gráfico de la intensidad de la luz en función de la longitud de onda. Cuando la célula con la muestra no está en su sitio, en el osciloscopio aparece una línea recta horizontal. En cambio, cuando la célula está colocada, la curva resultante identifica rápidamente las longitudes de onda a las que la muestra absorbió la luz.

Por lámpara, Morelli emplea una lámpara de coche tipo bulbo (como una bombilla ordinaria) de filamento recto. La lámpara se monta en un soporte orientado verticalmente. Aunque está calculada para trabajar a 12 volt, Morelli conecta la lámpara a 21 volt con el fin de obtener la suficiente intensidad y uniformidad a través de todo el espectro visible. (Evidentemente, el potencial mayor acorta la vida de la lámpara.)

La carcasa soporte de la lámpara está hecha con una lámina de acero muy delgada, doblada de forma adecuada. Morelli fabrica siempre la carcasa de dos piezas, de suerte que la lámpara pueda ser cambiada fácilmente. La parte posterior se une a la base principal del espectrofotómetro mediante un soporte en forma de *U*. La parte delantera incorpora un pequeño tubo, a cuyo través pasará la luz. En el extremo delantero de dicho tubo hay una placa plana con dos hojas de afeitar sujetas por cuatro tornillos. El tubo y las hojas de afeitar sirven para evitar que la luz difusa entre en el fotomultiplicador. Con este mismo fin hay otras pantallas colocadas dentro del aparato y éste debe mantenerse en el interior de una caja para impedir que le llegue luz de la habitación.

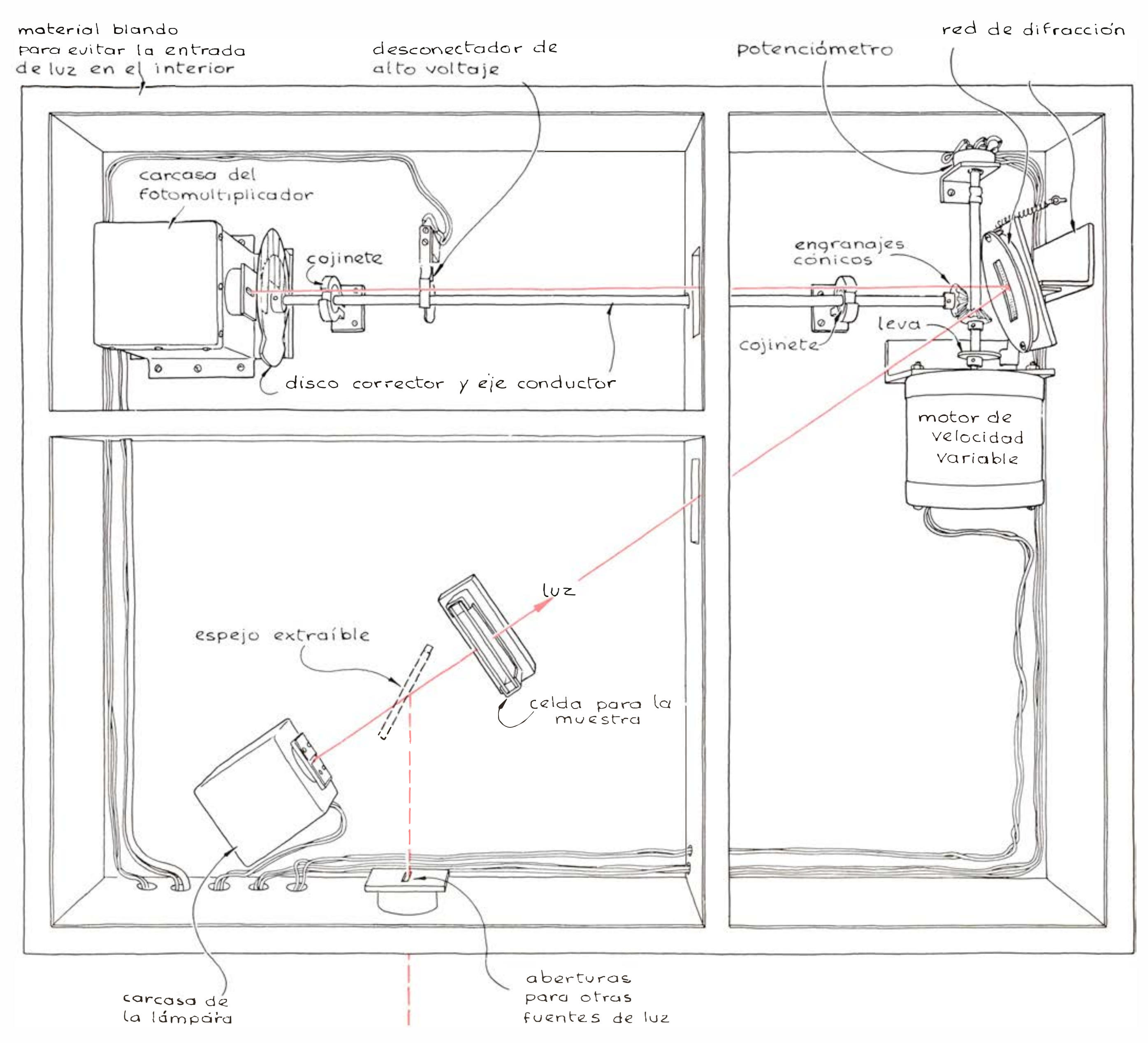
Morelli construyó la célula para la

muestra recortando una pieza en forma de *U* de plexiglás de medio milímetro, emparedándola con dos plaquitas de cristal y pegando las diversas piezas con un sellador de silicona. Como soporte de célula sugiere recortar y doblar una laminita de metal provista de una corredera y atornillarla a un taco de madera. La célula se coloca frente a la luz procedente de la lámpara. En el procedimiento normal de observación de las características de una solución se precisan dos celdas: para el disolvente, con el fin de establecer un patrón, y otra para la solución, que luego podremos comparar con ese patrón.

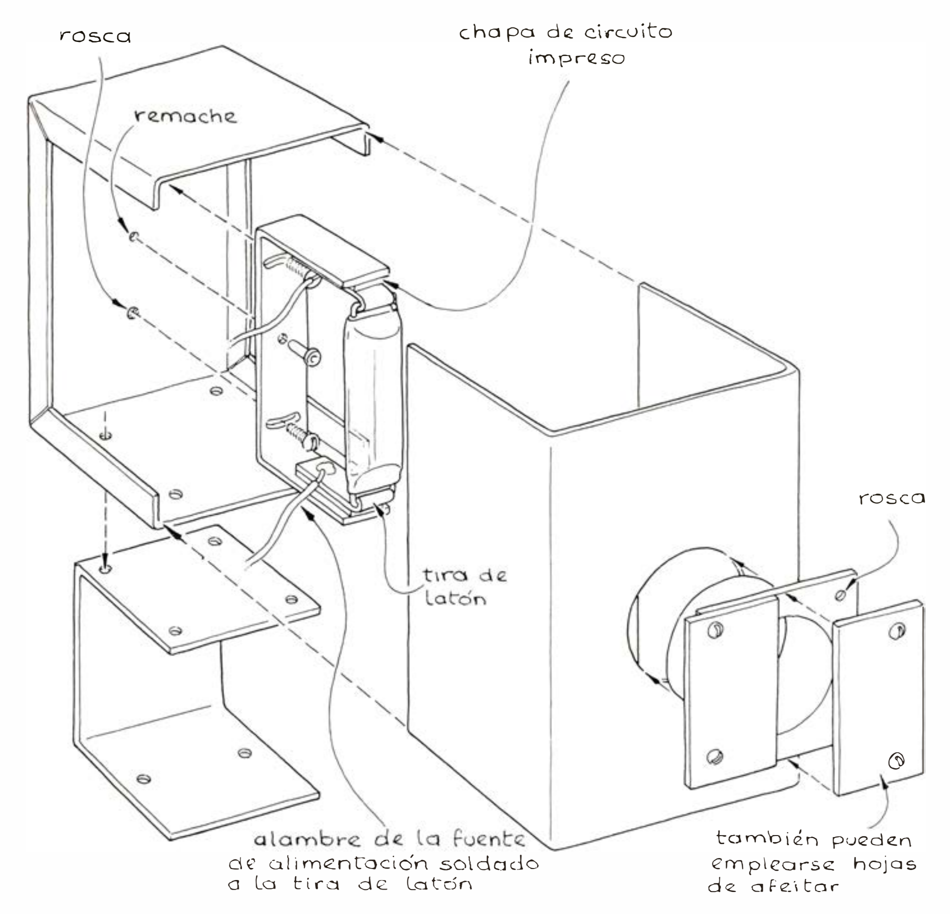
La red de difracción dispersa la luz transmitida a través de la célula con la muestra. Morelli emplea una red de difracción curva, del tipo de las que se encuentran en algunos espectroscopios astronómicos. (Antes de comprar tales redes se aconseja documentarse en las revistas de astronomía.) La red empleada por Morelli tiene un radio de curvatura de 50 centímetros, un área rayada de 3,175 por 4,445 centímetros y una densidad lineal de 1200 líneas por milímetro. Las líneas están situadas de manera que luz con una longitud de onda de 5000 angstrom tenga su máximo de reflexión en el primer orden de difracción. (En este caso, se dice que la red brilla a 5000 angstrom.) Se podrían utilizar otros tipos de redes, pero Morelli evita usar las que tienen un número *f* alto, pues ello comporta que el instrumento pierda sensibilidad. Si se usa una red plana, es necesario colocar una lente para poder proyectar la imagen de la abertura de la carcasa de la lámpara en la abertura del fotomultiplicador (por medio de una reflexión en la red).

La red de Morelli está montada sobre una placa de aluminio provista de una bisagra. De lo que se trata es de permitir que la red oscile horizontalmente alrededor de la bisagra y así barrer el espectro a través del fotomultiplicador. La bisagra está unida a un ángulo de aluminio que se encuentra pegado a la estructura principal del espectrofotómetro. La bisagra debe girar con suave firmeza, sin que llegue a oponer demasiada resistencia al movimiento. Morelli confiesa que probó con varias bisagras antes de encontrar la adecuada.

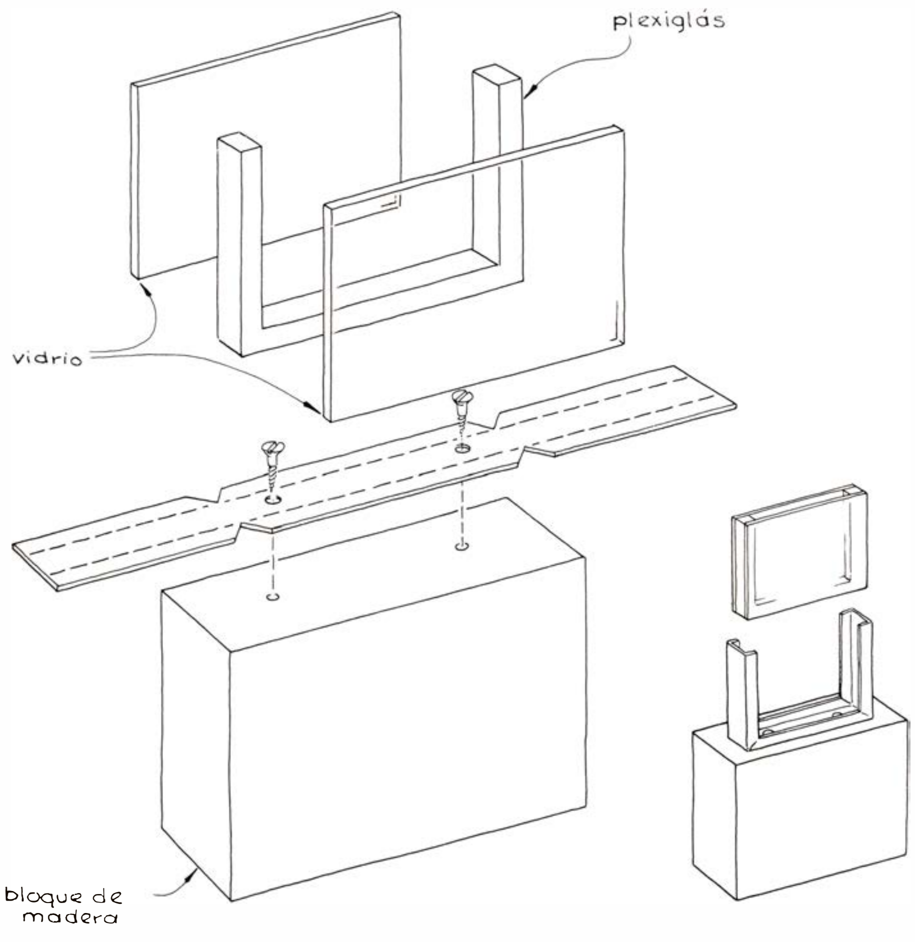
Para el ritmo de observación más alto, la red debería mantenerse de forma que oscilase alrededor del centro. Pero este punto de pivote no suministra el mejor enfoque para todas las longitudes de onda del espectro visible; la mejor distancia focal para una red depende de la longitud de onda. Para conseguir un buen montaje óptico, el punto de pivota-



Espectrofotómetro diseñado y construido por Dean Morelli



Carcasa de la lámpara.



Célula para la muestra con su soporte

ción debería estar desplazado unos seis centímetros del centro de la red. En ese punto, la velocidad de observación es más baja (debido a que la red tiene un momento de inercia mayor que el que tendría un pivote central), pero la mejora en el enfoque merece el sacrificio.

La red está hecha para moverse por medio de una leva que se apoya en un pequeño resalte situado en la placa sobre la que está montada la red. El resalte, en forma de ángulo recto, está sólidamente sujeto a la placa. Mantiene uno de sus extremos contra la leva por presión de un muelle que lleva incorporado la placa. Cuando la leva gira, el resalte se desliza a lo largo de su borde, y fuerza a que la red haga el recorrido horizontal. El muelle asegura que el resalte se mantenga pegado a la leva.

La leva es un recorte de una lámina de latón de dos centímetros. Primero se dibuja sobre un papel con su forma y tamaño reales. Se coloca luego el papel sobre el metal, se vuelve a repasar el dibujo, se recorta el metal por la marca impresa y, por último, se liman los bordes. Lo que se pretende es que la leva arrastre la red en un movimiento dentado, conduciéndola horizontalmente y de modo uniforme hasta alcanzar el final del barrido. Este movimiento se obtiene cuando el radio de la leva varía uniformemente con el ángulo alrededor de su centro. En dicho centro se practica un orificio, de unos 5 milímetros de diámetro, que ajuste con el eje del motor. Morelli soldó a la leva un aro de sujeción de latón procedente de un botón de un tablero de mandos. Esa sujeción solía quedar bajo la cubierta de plástico. El tornillo del aro de sujeción permitía a Morelli que ésta y la leva quedaran unidas con firmeza al eje del motor.

El barrido superior de la velocidad de barrido del espectrofotómetro lo fija la velocidad de oscilación de la red. Si la red va demasiado deprisa, el muelle no puede mantener el resalte contra la leva y el barrido se hace mal. Si se dispone un muelle más fuerte se puede distorsionar la velocidad del motor debido al rozamiento entre el resalte y la leva.

En el lado opuesto a la red y sobre la base del espectrofotómetro se encuentra el fotomultiplicador y su carcasa. Esta es metálica y de poco espesor. En su parte frontal se aplica un tubo metálico soldado o bien pegado con epoxi. Dentro del tubo va otro de diámetro menor. El tubo más pequeño lleva incorporadas unas cuchillas de afeitar de forma que dejen una leve abertura. Como el tubo pequeño es fácil de extraer, el experimentador puede instalar rápidamente aberturas de diferentes tamaños.

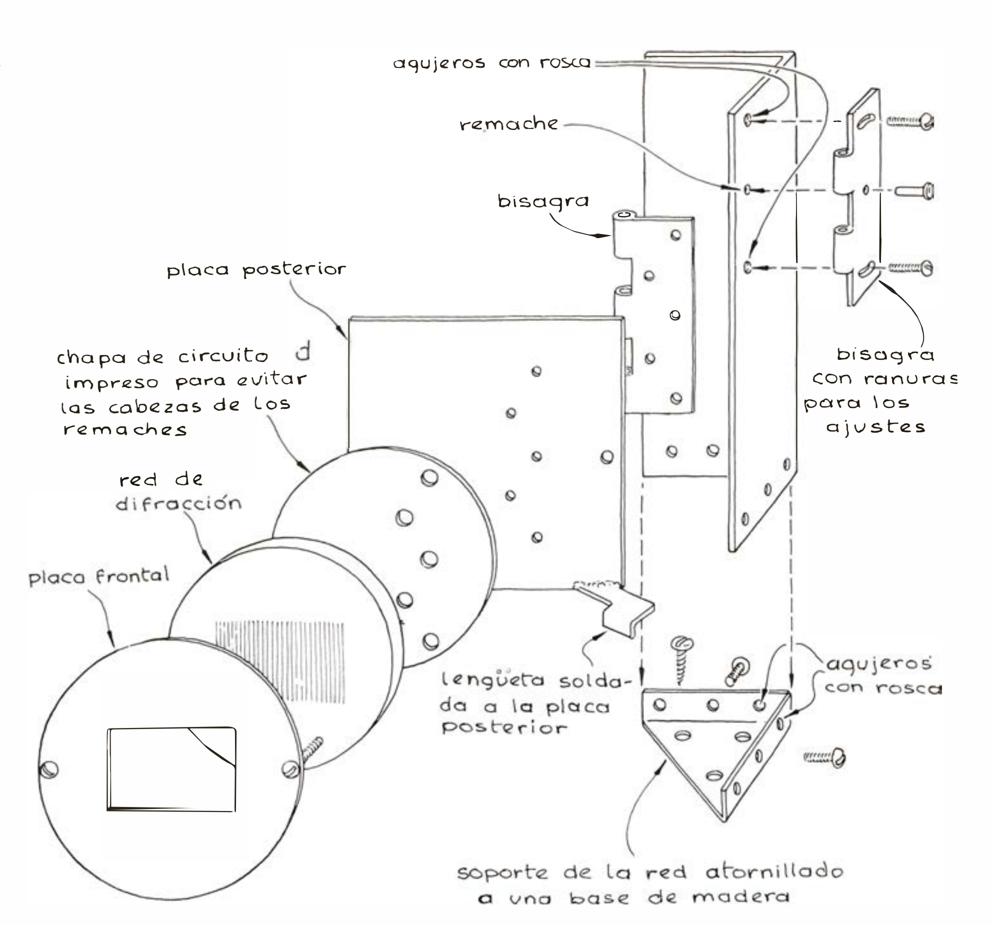
El tubo fotomultiplicador se monta sobre un soporte dentro de la carcasa. Las resistencias necesarias para suministrar la gama de potenciales adecuada al fotomultiplicador se sueldan directamente a éste a través de un agujero que tiene el soporte. Al igual que todas las demás parte metálicas del espectrofotómetro, la carcasa debe ir pintada de negro mate para absorber la luz difusa.

Morelli ha trabajado con varios tipos de tubo fotomultiplicador. El que uno puede comprar depende fundamentalmente de lo que encuentre en el mercado. El modelo 931-A, se consigue fácilmente. Morelli también usa el 1P21, que es un tubo del mismo tipo, sólo que tiene una eficiencia cuántica mayor y una corriente de fondo menor. La eficiencia cuántica indica cuántos electrones emitirá el fotocátodo del tubo por cada fotón que absorba. Si la relación es muy alta, se dice que es muy sensible y que tiene una eficiencia cuántica alta. Se entiende por corriente de fondo la corriente que pasa por el tubo cuando no está llegando luz al fotocátodo. Debe ser baja, pues así no enmascara la señal y en consecuencia no impide las mediciones de bajos niveles de luz. Tanto el tubo 931-A como el 1P21 tienen la superficie fotoeléctrica del tipo denominado S-4, compuesta de cesio y antimonio. Aunque los tubos muestran un rendimiento relativamente alto en la mayor parte del espectro visible, pierden sensibilidad hacia el final del rojo, con un límite en la práctica hacia los 6700 angstrom.

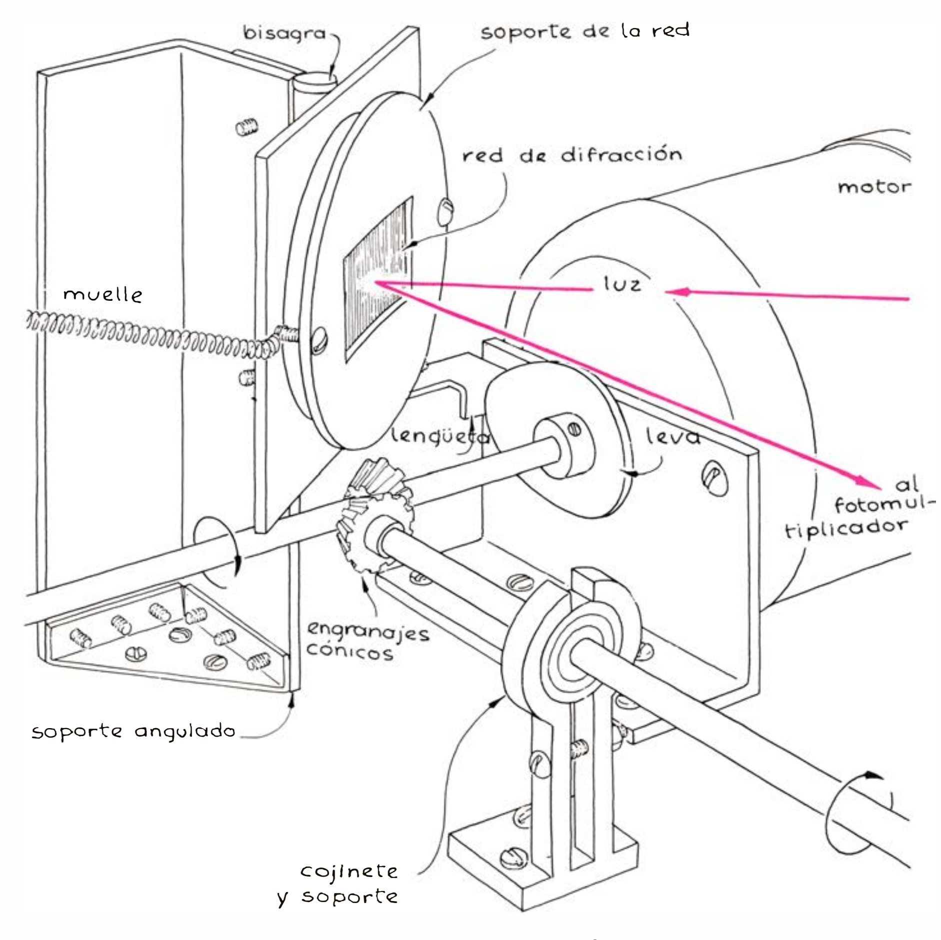
Hay que acertar a la hora de disponer el tubo fotomultiplicador, o lo que es lo mismo, éste tendrá una gama de potenciales adecuada. Morelli emplea una fuente de alimentación regulada que proporciona alrededor de 100 volt por etapa de cada una de las minietapas de su fotomultiplicador 931-A. Para información sobre diseños de fuentes de alimentación se puede consultar cualquier buen manual de radioaficionado.

Recientemente, Morelli probó un fotomultiplicador RCA 4840. Su superficie fotoeléctrica alcalina (llamada S-20) tenía un rango de espectro más ancho que el de su tubo 1P21, siendo sensible a longitudes de onda de hasta 8000 angstrom. El tubo tiene además una cubierta transparente a las longitudes de onda ultravioleta. Con este tubo Morelli logró deflecciones verticales completas en el osciloscopio midiendo la intensidad de la luz en un rango de 3400 a 7000 angstrom.

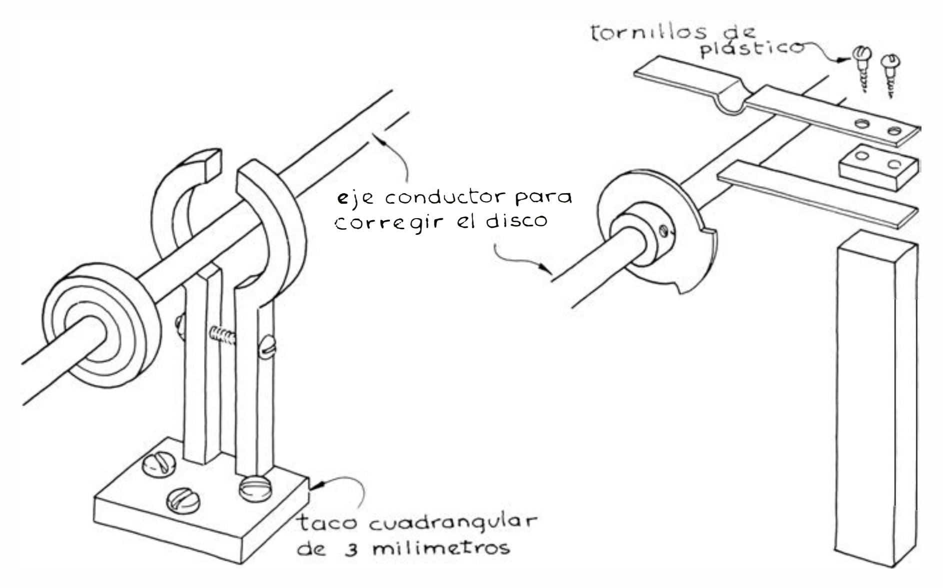
La salida del fotomultiplicador se conecta a la entrada vertical del osciloscopio, que debe tener un amplificador vertical de corriente continua. El brazo del



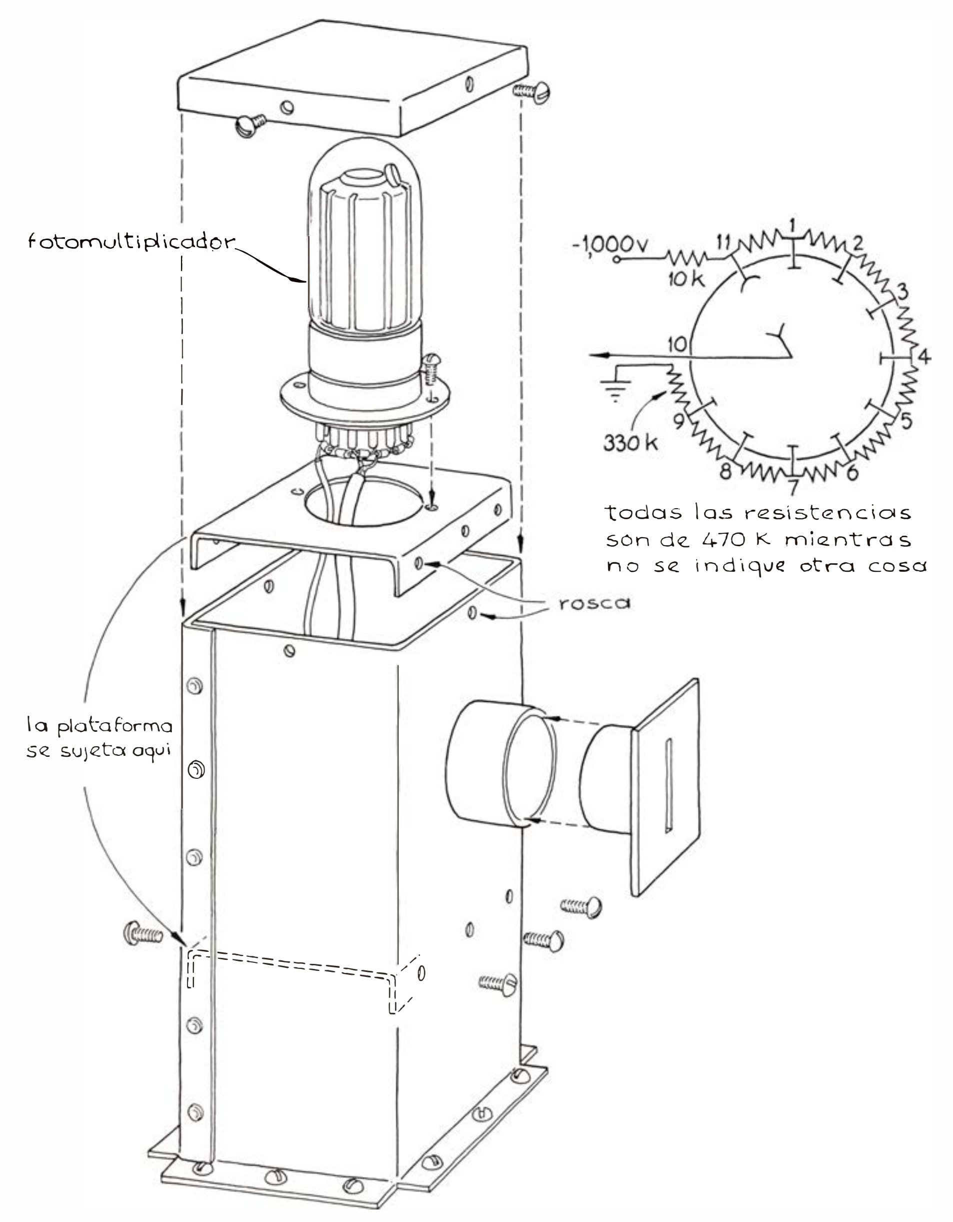
Detalles del montaje de la red de difracción



Modo de montar la red de difracción



Disposición del desconectador de alto voltaje



El fotomultiplicador y su carcasa

osciloscopio debe moverse a través de la pantalla de izquierda a derecha, conforme la red barre el espectro a través de la ranura de entrada del fotomultiplicador. La activación del osciloscopio, para comenzar cada barrido a través de la pantalla, depende del tipo de osciloscopio. Si está provisto internamente de la posibilidad de barrido provocado, la señal de disparo puede obtenerse por una pequeña luz, un disco con ranuras radiales y una fotocélula de silicio. El disco se monta en una prolongación del eje del motor que hace oscilar la retícula. Se sitúa la luz a un lado del disco y la fotocélula, al otro. Cuando la ranura del disco se posiciona correctamente, la luz de la bombilla pequeña incide en la fotocélula cada vez que la retícula comienza otro barrido del espectro a través del fotomultiplicador. La señal de la fotocélula dispara entonces el barrido del trazo del osciloscopio. (Morelli advierte que los osciloscopios baratos no "disparan" con fiabilidad a frecuencias de barrido bajas del espectrofotómetro.)

Se puede montar un sistema mejor para representar la salida del fotomultiplicador, cuando el osciloscopio posee un amplificador horizontal para corriente continua. Se monta un pequeño proyector en un tubo de metal de forma que proyecte la imagen de una ranura sobre una fotocélula de silicio (SM4) al otro lado de una leva torneada adecuadamente. La leva es similar a la que hace mover la red de difracción. Cuando la leva gira, provoca que poco a poco incida una luz sobre la fotocélula. La señal de la fotocélula se amplifica en la entrada horizontal del osciloscopio, mandando el trazo de izquierda a derecha a través de la pantalla. La segunda leva puede montarse en una prolongación del eje del motor de la red para sincronizar el barrido horizontal del osciloscopio con el movimiento de la red.

Si el amplificador de la entrada horizontal del osciloscopio no resultase suficientemente sensible para este sistema, el proyector y la fotocélula podrán sustituirse por un potenciómetro, que proporcionará una señal mayor a la entrada. En esta función, el potenciómetro debe girar continuamente, por lo que se eliminará el diente que impide que gire del todo. El potenciómetro se monta en el eje del motor con un trozo corto de tubo de goma. Una batería de 1,5 volt proporciona el voltaje a través del potenciómetro, cuya salida varía así continuamente de 0 a 1,5 volt, en diente de sierra, cuando el eje hace rotar la conexión interior del potenciómetro. La señal conducirá el trazo del osciloscopio suavemente a través de la pantalla.

Quizá prefiera el lector eliminar el

trazo de retorno del osciloscopio que se mueve a través de la pantalla de derecha a izquierda antes de cada nueva medida de intensidad. Para borrar el trazo de retorno, Morelli instaló un interruptor en el eje motriz, entre la montura de la red y la placa correctora del fotomultiplicador. El interruptor consta de dos láminas de metal, que se separan por una leva situada en el eje durante el período en que el trazo del osciloscopio deba eliminarse. El interruptor se conecta en serie con el hilo que suministra la alta tensión al osciloscopio. Pero puede ocurrir que nuestro experimentador quiera mantener el trazo de retorno del osciloscopio para que sirva como línea de referencia en la pantalla, indicando la línea base de intensidad de luz cero. Dicha línea base puede resultar de particular interés si el osciloscopio posee una fuente de alimentación no regulada, por lo que el trazo tiende a fluctuar.

Si el osciloscopio tiene baja sensibilidad en su entrada vertical, se necesita un preamplificador que aumente la señal del fotomultiplicador. Morelli utiliza amplificador operacional de circuito integrado UA741, de escaso coste, en un circuito realimentado simple, accionado con baterías (suficiente para dar ±15 volt.).

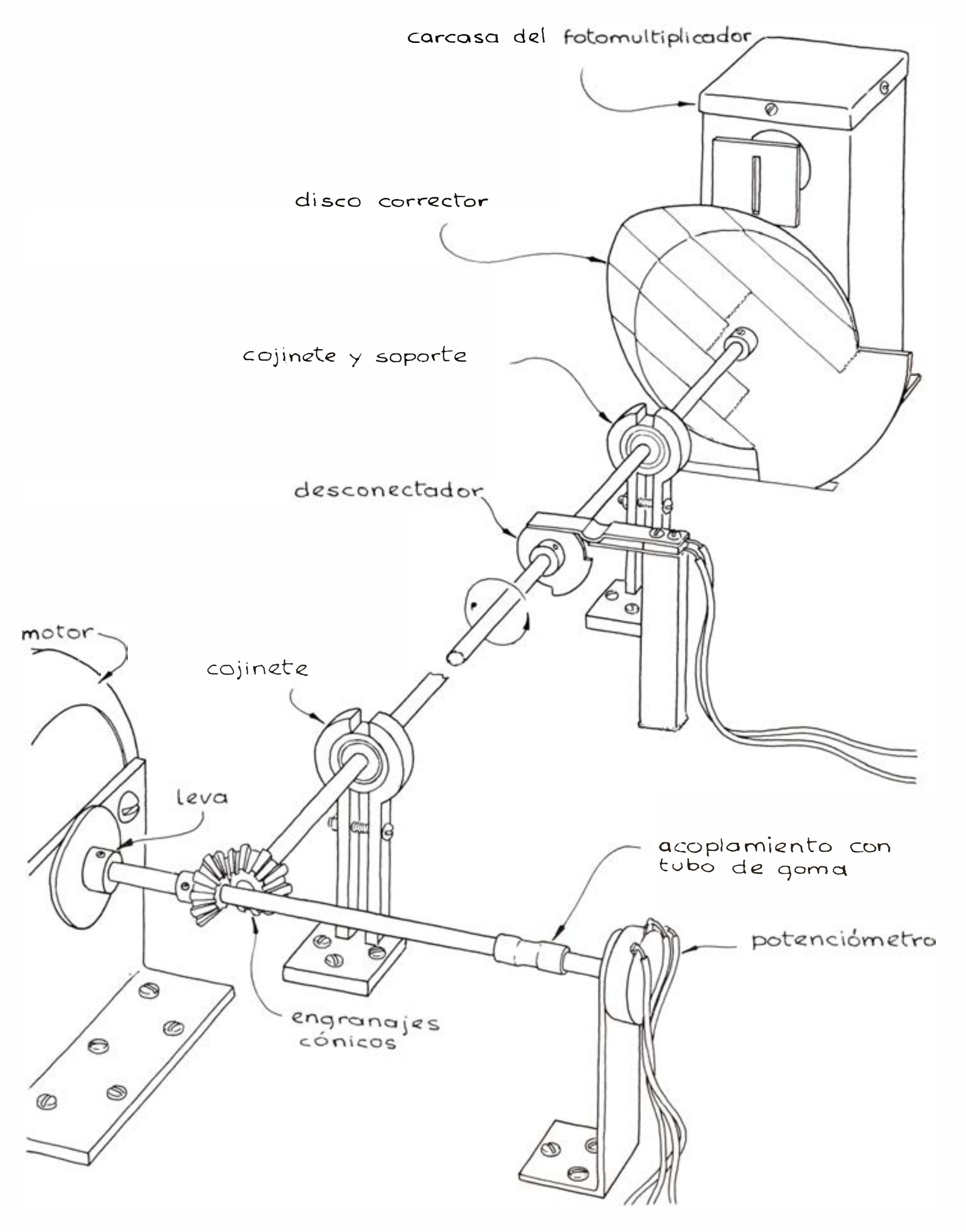
Morelli utiliza su instrumento como se indica a continuación, para determinar las características de absorción de una solución comparadas con las del disolvente puro. Primero consigue el trazo del osciloscopio con el disolvente puro en la célula de la muestra. Incluso en el caso en que el disolvente transmita perfectamente toda la luz, el trazo no será plano, ya que la intensidad de la luz emitida por la lámpara no se mantendrá constante a lo largo de todas las longitudes de onda del espectro visible. Tampoco la respuesta del fotomultiplicador será uniforme a través del espectro. El trazo del osciloscopio dibujará, pues, una línea más o menos acampanada, con su parte superior desapareciendo de la pantalla.

Para compensar estas faltas de uniformidad (y cualquier otra que distorsione la respuesta del instrumento a lo largo del espectro), se construye un disco corrector. Se monta el disco en un eje acoplado, mediante engranajes cónicos, al eje del motor que hace oscilar la red. El disco tiene por misión disminuir selectivamente la luz que incide en el fotomultiplicador a las longitudes de onda que dan un trazo más alto en el osciloscopio. Un disco corrector bien ajustado bloqueará la luz a cada longitud de onda del espectro, de suerte que el trazo sea aproximadamente horizontal en la parte alta de la pantalla. Luego, con un disolvente puro el lector obtendrá un trazo plano en el osciloscopio.

Morelli corta el disco corrector de una chapa de latón. Cuando se monta en el eje motriz, el disco bloquea parte de la luz que incide en la ranura frontal del fotomultiplicador. A continuación, coloca trozos de cinta aislante negra en el borde del disco, al tiempo que hace girar manualmente el eje del motor. Para cada posición del motor la red de difracción manda una determinada longitud de onda sobre la ranura; la desviación vertical del punto en pantalla del osciloscopio corresponde a la respuesta del fotomultiplicador a esa longitud de onda. A cada longitud de onda el punto está desviado inicialmente hacia fuera de la parte superior de la pantalla. Morelli añade un pequeño trozo de cinta adhesiva al disco corrector al objeto de bloquear la luz lo suficiente como para que el punto baje hasta la parte superior de la pantalla. Entonces hace girar suavemente el eje del motor y repite el proceso. Dice que, aunque el sistema es algo tosco, el trazo que resulta en la parte superior de la pantalla es plano con un error de un 5 por ciento.

Cuando la célula que contiene sólo disolvente se sustituye por otra con una sustancia disuelta en ese disolvente, la absorción de la luz por la solución a ciertas longitudes de onda desencadena picos que se extienden hacia la parte inferior de la pantalla del osciloscopio. Si la disolución lograra absorber totalmente la luz a alguna longitud de onda, el pico correspondiente alcanzaría el borde inferior de la pantalla.

La resolución del espectómetro depende de la calidad de la red y de la anchura de la ranura situada delante del fotomultiplicador. Una ranura más estrecha toma una banda más estrecha de longitudes de onda del espectro que la



Montaje del mecanismo de barrido

red lanza a su través. En la posición de la ranura, la anchura del espectro visible completo es de 125 milímetros. Morelli emplea una ranura de unos cuatro milímetros de ancho, que admite por tanto 4/125 (alrededor de 120 angstrom) del espectro completo del multiplicador. El lector puede aumentar la resolución cambiando la ranura de cuatro milímetros por otra más estrecha; en contrapartida, perderá intensidad de luz. Entonces, el fotomultiplicador responderá sólo a la luz de longitud de onda de su rango más sensible (para el tubo 931-A de Morelli, aproximadamente entre 3700 y 6000 angstrom) y se perderá la parte roja del espectro. Con una ranura de cuatro milímetros puede observarse un rango más ancho (3500-6300 angstrom). El límite inferior del rango viene impuesto entonces por la absorción ultravioleta del vidrio de la bombilla, el tubo fotomultiplicador y la célula de la muestra. Otro inconveniente que resulta de las ranuras inferiores a cuatro milímetros de anchura es que quizá necesitarían un movimiento excéntrico de la retícula para evitar el desenfoque en los extremos del espectro.

Morelli calibra su espectómetro poniendo un tubo Geissler de emisión de hidrógeno, en vez de la lámpara acostumbrada. Varias de las líneas del espectro visible emitidas por el hidrógeno (la serie de Balmer) pueden identificarse en la pantalla del osciloscopio. Cuando se pone una ranura de cuatro milímetros las líneas aparecen como picos agudos. Como las longitudes de onda de tales líneas se conocen exactamente, Morelli puede calibrar el eje horizontal del trazo en pantalla, en angstrom.

De intención no me he entretenido en la precisión de las dimensiones del espectofotómetro; Morelli insiste en la necesidad de modificar y ajustar, por parte del experimentador, el diseño básico para adecuarlo al tipo de equipo que haya sido capaz de reunir para construir el instrumento. Los alojamientos del fotomultiplicador y de la lámpara deben estar tan cerca como sea posible. Morelli sugiere una forma de determinar las distancias adecuadas entre la lámpara, el fotomultiplicador y la red. Colocar en una mesa un tubo Geissler, la red y una ranura. Con las luces de la habitación apagadas y el tubo encendido desplazar los tres objetos por la mesa hasta que el espectro quede bien enfocado sobre la ranura. Las distancias así determinadas pueden servir de guía para la construcción del espectrofotómetro.

A la hora de recurrir al espectrofotómetro para examinar las características del espectro de fuentes externas de luz, como lámparas o llamas, se monta un espejo en el instrumento delante de la lámpara. La luz de la fuente externa entra a través de un agujero a un lado de la caja que aloja el espectrofotómetro y se dirige hacia la red mediante el espejo. La fuente ha de ser relativamente brillante para que no se pierda el extremo rojo del espectro por causa de baja sensibilidad del fotomultiplicador en este extremo.

Las aplicaciones posibles al espectrofotómetro de barrido rápido de Morelli son numerosas. Además de las reacciones químicas oscilantes que ya hemos mencionado, hay otros fenómenos ópticos, casi innumerables, que cambian con excesiva rapidez para cualquier espectofotómetro que necesite examinar, a base de mediciones punto por punto a través del espectro visible. Morelli, por ejemplo, pretende examinar la fosforescencia de algunos cristales.

Algunos astrónomos aficionados podrían mostrarse interesados en analizar la composición del espectro de la luz solar. Morelli utiliza un telescopio y un espejo para proyectar la imagen del sol en un agujero de alfiler instalado provisionalmente en lugar de la ranura lateral de la caja que aloja el espectrofotómetro. Mueve la imagen del sol a través del agujero, y así consigue registros en su osciloscopio de espectros de distintas zonas del sol. Podemos sacar observaciones similares con otros objetos celestes, pero debido a la menor intensidad de su luz, el espectrofotómetro con el modelo básico de Morelli respondería principalmente a las longitudes de onda a las cuales el fotomultiplicador se muestra más sensible.

Morelli no ha dejado de trabajar sobre su instrumento. Su próximo gran cambio será reemplazar el sistema mecánico de registro del diseño actual por otro totalmente electrónico, basado en un vidicón (un tubo almacenador de imágenes). Este sistema será más sensible porque la señal del osciloscopio estará formada por el almacenamiento de información de muchos registros. Sería, por ello, mejor para el análisis espectral de la luz procedente de fuentes débiles, tales como cristales fosforescentes. Morelli podrá examinar fenómenos ultrarrápidos: el fogonazo de una lámpara de flash de alta velocidad. El vidicón es más sensible en el extremo rojo del espectro, de manera que el instrumento no tendrá uno de los defectos de los basados en el fotomultiplicador. Morelli está construyendo ahora un espectofotómetro de registro rápido con un tubo de vidicon de los que se usan en circuitos cerrados de televisión en blanco y negro.

Libros

Tres frentes de interés de la biología contemporánea: paleontología, información biomolecular y evolución

Emiliano Aguirre, Antonio Fontdevila y Luis Alonso

PALEONTOLOGÍA. Tomo 2. Vertebrados: Peces, Anfibios, Reptiles y Aves, por B. Meléndez. Paraninfo, Madrid 1979. No abundan los libros de paleontología originales en español. No es un tema popular ni divulgado aquí. Y, sin embargo, España es un país rico en fósiles, hasta el punto de poder afirmarse que los yacimientos paleontológicos constituyen uno de los recursos culturales y científicos más notables de estas tierras: pero recurso inexplotado, al menos para las instituciones y lectores de habla hispana. Otro tanto se puede decir de varios países hispanófonos del Nuevo Mundo. Normalmente, la literatura y las imágenes de dinosaurios y otros fósiles nos llegan de distintas provincias culturales; aun así, tampoco se traduce mucho ni lo mejor. La literatura paleontológica en general es una de las lagunas más importantes y sensibles de la actual cultura hispánica, uno de los índices más claros de su descompensación, si no retraso.

Más triste aún es, y causa de la anterior, al menos en parte, la falta de interés por la paleontología en los mismos medios científicos, de investigación y profesionales. Parecen avergonzarse de la ciencia paleontológica las mismas instituciones oficiales españolas: en vano buscará alguien, docto o simple turista, un Instituto de Paleontología o un Museo de Paleontología en España, fuera de Sabadell, donde costó años establecer y a duras penas se mantiene uno provincial, de primera línea, o del modesto Museo paleontológico "in situ" de Ambrona, no amparado por institución alguna, carente de dirección y de presupuesto. La docencia de la paleontología no parece calar hondo en los mismos profesionales que la necesitan ni ser justamente atendida en los claustros universitarios ni en los antros ministeriales o rectorales. Por eso no puede culparse sólo a los editores de este hiato cultural. No obstante, la gente desea saber, pregunta, muestra curiosidad; si bien, por falta de puntos de referencia, su receptividad es como enclenque o desorientada.

Hacen al caso estas tristes consideraciones, porque el principal defecto del tratado de paleontología de Bermudo Meléndez es externo a la obra y al autor: es su misma singularidad y desamparo. Si bien merece un sonoro pláceme y bienvenida esta obra, y una felicitación la editora que se atreve a publicarla en tan adverso ambiente, aquella viene a resultar un libro de texto sin contexto.

Este volumen forma parte de un tratado, "Paleontología", cuyo desarrollo prevé el autor en cuatro tomos: éste es el segundo en orden didáctico y también el segundo en publicarse, después del tomo primero, "Parte general e Invertebrados", aparecido en 1970 (que incluye también todos los microorganismos fósiles, tratados con excesiva concisión). Dentro del plan general de la obra, el autor principal justifica adecuadamente en el prólogo el dedicar un tomo a los Vertebrados, con exclusión de los Mamíferos, y la extensión que se concede a los diversos grupos es proporcionada. También es equilibrada la distribución de la materia en dieciséis capítulos, la mitad casi de los cuales corresponde a ramas reptilianas, por el esplendor evolutivo y el papel en los paleoambientes de esta fecunda clase.

El carácter de libro de texto universitario, de paleontología sistemática, diría que "clásica", y el empeño didáctico, son las cualidades que más confieren a la unidad de la obra. Los modelos serían el célebre tratado de Romer, sobre paleontología de Vertebrados, en un tomo, y el más amplio "Traité de Paléontologie" de Piveteau, en varios, verdadera obra enciclopédica. Meléndez logra un contenido, en parte más completo que el primero y más sintético que el segundo, a la vez que más didáctico, con un loable esfuerzo de síntesis, que sólo decae en los capítulos de los Sinápsidos. El desarrollo de la paleontología, por otra parte, hace imposible una obra, incluso de la sencillez de este tomo, sin la colaboración de varios autores. La realización de Bermudo Meléndez es pionera en la ciencia española no sólo por el libro mismo y su contenido, sino por haber formado el equipo de jóvenes paleontólogos españoles capaces de tratar con maestría y en profundidad estos temas, y varios de ellos, si no todos, muy al corriente en la bibliografía y cuestiones controvertidas.

Las desigualdades que se aprecian en el tratamiento de los diversos capítulos puede atribuirse a la concurrencia de varias condiciones. En unos casos nos encontramos con fósiles de mayor complejidad morfológica que en otros; hay grupos con numerosísimas especies y familias, como los Teleósteos, pero con menor interés biostratigráfico y evolutivo que otros no tan abundantes pero más problemáticos. Es lógico advertir más originalidad, más dominio y fluidez en cuestiones controvertidas y de síntesis, en los capítulos escritos por especialistas, como los de Anfibios, Sinaptosaurios, Dinosaurios y Quelonios, que enriquecen con meritorios apartados de ecología y etología. Por otra parte, encontramos menciones de fósiles representativos de yacimientos españoles, lo que parece justo, a menudo ilustrados con fotografías o dibujos de conjunto, pero sólo son excepcionales las ilustraciones de análisis anatómico basadas en estos fósiles españoles: éstas, en general, se copian de tratados extranjeros; un indicador más, a fin de cuentas, de la falta de una tradición de anatomía comparada en nuestro país y área cultural, y de paleontólogos vertebristas con tiempo y medios para desarrollar esta complicada labor. También se acusa la dependencia de fuentes francesas y anglosajonas; sólo excepcionalmente o muy de pasada se tratan los fósiles sudamericanos, casi siempre a través de fuentes extrañas, sobre todo en los Terápsidos; los autores hispanoamericanos están ausentes.

Salvo pequeñas lagunas, como los Traversodóntidos, que se resuelven en seis líneas muy vagas (p. 473), cuando parecen tener gran importancia en la investigación del origen de los verdaderos Mamíferos (y dentro de un capítulo que más bien peca de nimio en multitud de familias e incluso a nivel de géneros), la información taxonómica es muy completa; pero se echan de menos en otros grupos, también de interés, cuadros sinópticos de distribución biostratigráfica y geográfica, como los que ilustran, por ejemplo, los Conodontos y parte de los Sinápsidos (aunque algunos de estos últimos están algo atrasados). Los datos paleogeográficos, conocidos o problemáticos, que mucho tienen que ver con los tanteos y logros evolutivos de los Vertebrados, faltan casi por completo. En general, el esfuerzo para conseguir una

ilustración tan abundante como ecléctica es uno de los logros indiscutibles de la obra.

Entre las cuestiones más importantes, para seguir el hilo conductor de la historia de la vida, se cuentan los parentescos entre los grandes grupos de Vertebrados, las clases y las subclases. En el libro se encuentra bien tratado el problema del origen de los Vertebrados, el de los Anfibios, los diferentes linajes de dinosaurios y las Aves. Algo confusa queda, en cambio, la cuestión –por lo demás controvertida aún– de los primeros peces con mandíbulas, esto es los posibles parentescos de los Acantodios, que se incluyen, sin justificación suficiente, con los Placodermos, en un gran título, "Paleictios"; este nombre no corresponde a una nomenclatura aceptable como los que encabezan otros capítulos, aun cuando lo sea su derivación etimológica. Los posibles entronques de los grupos más importantes de Reptiles, Arcosaurios, Lepidosaurios y Sinápsidos, con remotas formas reptilianas carboníferas, y los posibles parentescos entre ellos, no rebasan actualmente el estado de débiles hipótesis, pero se echa de menos una mención, siquiera de pasada. La posición de los *Seymouriamorpha* en la figura 32 (p. 60) contradice a lo que se afirma en el texto sobre este grupo y el origen de los Reptiles; tal vez no haya sido completa la crítica y el cuidado en el uso de las fuentes, que en ciertas cuestiones pueden ser obsoletas. Menos aceptable es que el término "Dinosaurios" encabece un capítulo: incluso se introduce subrepticiamente este término entre los órdenes del cuadro de la página 219, aun cuando en el texto se insiste en que carece de valor taxonómico, sino sólo vernáculo y popular, dada la independencia filogenética de los dos órdenes que en él se agrupan: no tratándose de un libro popular, sino riguroso en toda su extensión, no hay motivo para esta concesión. También choca el barullo de citas de Tanganyka, Tanganica, Tang. y Tanz. (pp. 470, 452, 419), tratándose en dos casos de cuadros "originales": la antigua colonia es hoy el estado soberano de Tanzania. Entre los defectos menores, ocurren ciertas expresiones ingenuas o inadecuadas, que pueden herir la sensibilidad de muchos biólogos, como cuando se habla de "perfeccionamientos", de animales "grotescos" y de "máxima autoridad". Y, aun cuando se trate de una cuestión puramente formal, empaña el mérito general de la obra una falta de criterio para la transcripción vernácula de muchos nombres latinos que salpica todo el tratado con grafías híbridas, como "Ichthyosaurios", "Captorhinomorfos", "Chelyidos", cuando el Diccionario de la Real Academia de la Lengua deriva de las mismas raíces las palabras Ictiosauro, Ictiología, Platirrinos, y en el mismo libro se escribe "Quelonios": en otra edición debería hacerse un esfuerzo para subsanar este defecto.

No hay que olvidar, en ningún aspecto, que este volumen, por su seriedad y por su singularidad, sentará cátedra y establecerá un magisterio en las universidades de habla hispana. Además, la presencia de la *Paleontología* de Meléndez, y en especial de este segundo tomo, en las librerías y centros de enseñanza, debe abrir camino al interés por estos temas en diversos niveles culturales, y preceder la aparición de nuevos libros sobre los fósiles y sobre la increíble dinámica de la vida, de la que ellos son testimonio, muchas veces curioso y siempre veraz. (E. A.)

ACERCA DE LA EVOLUCION, por J. Maynard Smith y H. Bume Ediciones, Madrid 1979. Estoy convencido de que la mayor parte de los biólogos actuales consultados individualmente no se atreverían a sostener que la formación evolutiva es secundaria para un biólogo. Y sin embargo, ¿por qué la mayoría de nuestros biólogos no han oído hablar apenas de evolución durante su formación académica? Afortunadamente algunos cursos de genética tienen en nuestras universidades una orientación evolutiva importante, pero los biólogos adquieren la mayor parte de sus conocimientos evolutivos en lecturas al margen de sus estudios universitarios. Estas lecturas marginales son de gran importancia no sólo para la formación biológica de nuestros especialistas, sino también para la cultura biológica del país.

Desgraciadamente muchas obras que sobre evolución biológica aparecen publicadas en nuestro país no reflejan el momento actual del conocimiento evolutivo. Y esto se debe básicamente a que son obsoletas y/o están escritas por autores no evolucionistas. Las razones de esto son, en el primer caso, el retraso tradicional con que llegan a la lengua española los conocimientos científicos y, en el segundo caso, la también tradicional y errónea costumbre de que la evolución biológica es un tema sujeto a la especulación literaria. No es este el caso del libro titulado "Acerca de la evolución" que comentamos. Esta obra recoge un conjunto de artículos aparecidos originalmente en su mayor parte en revistas especializadas que abarcan desde el año 1952, el más antiguo, hasta el año 1976, el más moderno. Su autor, John Maynard Smith, es una de las primeras personalidades evolucionistas actuales. "Acerca de la evolución" contiene la traducción de la obra original "On Evolution", publicada en 1972, más el artículo "Etica y evolución humana", incluido al final del libro.

Los artículos abarcan un amplio espectro de temas evolutivos, todos ellos de interés, que van desde problemas básicos (la selección de grupo, la preadaptación, los modelos del desarrollo, el polimorfismo, el sexo) hasta problemas generales (el neodarwinismo actual, el progreso evolutivo) y de influencia sociopolítica (La evolución y la historia, La eugenesia, la ética y la evolución). A pesar de la diversidad temática, existe una continuidad subyacente que discurre por todos los artículos y que está basada en la preocupación del autor sobre el estado actual del neodarwinismo como hipótesis científica. Maynard Smith trata de fundamentar con acierto que el neodarwinismo no es una teoría tautológica y es capaz de ser rebatida. En el capítulo "La situación actual del neodarwinismo", que se publicó por primera vez en 1969, el autor aborda abiertamente el problema planteando situaciones "absurdas" que no podría explicar el neodarwinismo. Algunos lectores podrán echar en falta en este artículo alguna referencia a la llamada (mal) evolución no-darwiniana. Es difícil saber si Maynard Smith omitió hablar aquí de este tema, porque entonces todavía no se había desencadenado la controversia neutralista en todo su esplendor o porque consideró, contrariamente a la opinión de algunos, que la teoría neutralista no altera básicamente el mecanismo neodarwiniano, sino que más bien excluye del mismo a la mayoría de las sustituciones aminoacídicas observables en las proteínas. En el capítulo "Las causas del polimorfismo", Maynard Smith hace un análisis fino de la teoría neutralista aplicada a la hemoglobina humana. La conclusión de que dicha molécula no ha experimentado prácticamente mutaciones neutras, a no ser que el tamaño de la población humana haya pasado por un "cuello de botella" en el último millón de años, constituye un ejemplo claro de que el conocimiento de parámetros ecológicos como el tamaño de población es básico para comprobar hipótesis neutralistas. La preocupación por una síntesis ecológico-evolutiva está presente en la mayoría de esta colección de artículos. En opinión de Maynard Smith la ausencia de teorías ecológicas adecuadas es uno de los fallos del neodarwinismo (p. 92). Las discusiones sobre la explicación neodarwiniana del aumento de comple-

jidad son estimulantes y plantean uno de los problemas actuales en la biología de sistemas. R. Margalef ("Perspectivas de la teoría ecológica". Blume. 1978) comparte esta preocupación en sus obras y trata de enfocar el problema por el lado del aumento de información utilizando entidades abstractas (diversidad, relación biomasa/productividad, etc.) que sirvan de índices evolutivos. En el capítulo: "El tiempo en el proceso evolutivo", Maynard Smith discute la validez del teorema fundamental de la selección natural, formulado por Fisher, para justificar el "progreso" o aumento de complejidad evolutiva. Después de un análisis brillante, el autor llega a la conclusión de que este teorema no permite explicar la evolución direccional de la complejidad. El problema fundamental radica en que no es posible comparar la adaptación entre dos unidades biológicas (poblaciones, ecosistemas, etc.) aisladas en el tiempo o en el espacio, dado que la adaptación es un concepto relativo y no absoluto. En este sentido, la adaptación fisheriana no es un parámetro absoluto como la entropía (información) que permita distinguir los estados sucesivos. La sugerencia de Margalef de que la sucesión ecológica tiene ingredientes que la asimilan a un proceso evolutivo podría arrojar luz a este problema.

Además de los aspectos de fondo discutidos en los párrafos anteriores, que por sí solos justifican ya la lectura de esta colección de artículos, abundan los temas con "garra". En el primer artículo, "Teoría de juegos y evolución de la lucha", el autor trata de justificar que la selección individual puede explicar la evolución de una estrategia de grupo. Para este desarrollo utiliza la analogía de la teoría de juegos de un modo brillante. Es posible que la idea de estrategia evolutiva estable no satisfaga a muchos como estrategia individual porque consideren que es una estrategia de grupo que aumenta la probabilidad de supervivencia del mismo, pero el desarrollo del tema en términos analógicos es de gran valor científico.

El tratamiento de la evolución del vuelo (La importancia del sistema nervioso en la evolución del vuelo) es corto pero delicioso. En este artículo Maynard Smith muestra su antigua formación técnica hablando de las condiciones de estabilidad del vuelo. El artículo es superficial, las referencias son antiguas, pero probablemente la complejidad del tema y la falta de estudios sobre el mismo justifican este tratamiento. Hay un error grave en la traducción del resumen y conclusiones. En las dos últimas líneas de la página 39 debe decir: ..."la evolución del sistema sensorial y nervioso hizo que la estabilidad que se encuentra en las formas primitivas ya no fuera necesaria".

Otros aspectos evolutivos que merecen la atención de Maynard Smith son los problemas epigenéticos y el significado del sexo. En el primer caso (La disposición de las cerdas en *Drosophila*) el autor describe un modelo de desarrollo apoyándose en sus numerosos trabajos sobre evolución de la disposición de quetas en *Drosophila*. Los avances en el campo de la genética del desarrollo han sido importantes desde la publicación de este artículo en 1962 y por este motivo únicamente se echa a faltar una interpretación más actual del problema. Sin embargo, algunas ideas siguen siendo muy valiosas y originales y en conjunto el artículo tiene valor como precursor de ideas en un momento en que todavía falta mucho por demostrar en el campo epigenético. El artículo "El origen y mantenimiento del sexo" plantea nuevamente la disyuntiva de la evolución del sexo como producto de selección de grupo o de selección individual. Maynard Smith trata de concretar en qué situaciones el sexo es ventajoso llegando a la conclusión de que "los metazoos deberían dar lugar con mucha frecuencia a variedades partenogenéticas". La ventaja evolutiva a largo plazo del sexo se explica más fácilmente mediante una selección de grupo, a la que Maynard Smith opone la dificultad de que si a corto plazo el sexo es menos eficaz que la reproducción asexual, la selección individual actuará en contra de la aparición del sexo. Posiblemente la explicación evolutiva del sexo seguirá como tema de controversia durante mucho tiempo, sobre todo si se sigue enfocando el problema evolutivo como simple proceso de sustituciones génicas y no como un aumento de nuevas funciones debido a la combinación de productos genéticos nuevos que muchas veces aparecen como sucesos únicos en forma de cambios cromosómicos.

He dejado para el final el comentario de los artículos de interés sociológico porque, en cierto modo, forman una unidad dentro del libro. La publicación en 1974 del libro *Sociobiology*, escrito por E. O. Wilson, sirvió para actualizar el viejo problema de la relación entre la evolución biológica y la evolución cultural. En el artículo Etica y evolución humana, Maynard Smith hace una amplia referencia a esta obra, comparando su alcance con las ideas de Waddington. En este y en los demás artículos sociológicos (Evolución e historia; Eugenesia y utopía) básicamente se plantean las siguientes cuestiones. ¿Es posible que las ciencias sociales permitan descubrir procesos que sean aplicables a las ciencias naturales, formando con ellas un todo? ¿Hasta qué punto el comportamiento social humano es explicable mediante los mismos mecanismos que rigen en las comunidades animales? Por lo que respecta al primer punto, Maynard Smith utiliza la analogía entre historia y evolución como una posible fuente de inspiración para sugerir mecanismos evolutivos, pero nunca como una prueba "a priori" de dichos mecanismos, las cuales deben comprobarse experimentalmente. De este modo, sale al paso de la errónea utilización de la dialéctica histórica como justificante de teorías de base lamarckiana. Tampoco es lícito utilizar la ética naturalista, basada en que lo "natural" es lo bueno, para defender posturas sociales. Como ejemplo de esto Maynard Smith utiliza el argumento feminista. El conocimiento de la evolución biológica del hombre podría, en todo caso, explicar el dominio del macho sobre la hembra, pero en el contexto social actual está muy lejos de justificar el "machismo". Las críticas de Maynard Smith al libro de Wilson son tremendamente constructivas y en este sentido, se opone a la crítica que el grupo de Boston, con Richard Lewontin a la cabeza, hizo en un tono poco racional del problema. Suscribo bastante la postura de Maynard Smith porque creo que el determinismo biológico propuesto por Wilson no es independiente del ambiente, sino más bien que las respuestas posibles del hombre ante un ambiente se comprenden mejor teniendo en cuenta su naturaleza biológica. Es una lástima que este último capítulo del libro no contenga una lista de referencias para el lector interesado.

El capítulo Eugenesia y utopía es en mi opinión el menos interesante por dos razones. En primer lugar, al haberse escrito en 1965 está obsoleto y, en segundo lugar, no aporta muchas novedades a lo que otros autores anteriores (ver referencias) habían dicho. Es preciso advertir que el término ingeniería biológica utilizado por Maynard Smith no tiene nada que ver con la acepción actual de ingeniería genética. El primero es un término fisiológico, mientras que este último se refiere a cambios dirigidos en los genes. Dos aspectos interesantes del artículo: el énfasis puesto en el coste social de la acción médica actuando en contra de la selección natural al aumentar la frecuencia de genes deletéreos en las poblaciones y la prioridad por el control de la natalidad frente a las prácticas eugenésicas cuya aplicación y resultados son más conflictivos. Ambos puntos de vista han sido avanzados hace tiempo por muchos autores, pero conviene re-

cordarlos de vez en cuando sobre todo en determinados contextos sociales un poco de espaldas a la evolución biológica.

Aparte de ciertos detalles originales que adolecen de falta de claridad (por ejemplo, la figura 3, página 46) o de errores tipográficos (pág. 111: I = 0.96 en vez de I = 0.06) y de algunos errores de traducción (hospedador parece una mejor traducción de "host" que huésped; herencia *no*-mendeliana mejor que nada mendeliana) y de impresión, la obra está bien presentada y su lectura resulta agradable. Es una suerte poder contar con esta colección de artículos sobre la evolución en lengua española y, sin duda, su lectura es recomendable, por no decir necesaria, no sólo para los estudiosos de la biología, sino también para cualquier persona que desea alcanzar un nivel cultural propio del tiempo actual. Incluso los especialistas que ya conozcan al menos parte de la obra encontrarán un placer estimulante al volver a ella. Yo por lo menos lo he sentido. (A. F.)

BIOMOLECULAR INFORMATION THEORY, por Serafín Fraga, K. M. S. Saxena y Manuel Torres. Ed. Elsevier Scientific Publishing Company, Amsterdam-Oxford-New York, 1978. La historia reciente de la biología es un avance hacia las últimas unidades de la vida. Como la historia reciente de la física ha supuesto un progreso espectacular hacia la búsqueda de las últimas partículas fundamentales: átomos, hadrones, leptones, quarks y gluones. Y al igual que en el terreno de la física la experimentación y observación no llega a confirmar o negar los postulados teóricos, la biología roza también la frontera de la especulación en las postreras consecuencias de la bioquímica cuántica.

¿En qué se distingue la biología de la información? ¿No es acaso un código –un conjunto de señales ordenadas– el fundante de la vida? Toda la bioquímica está atravesada de una jerga lingüística o informativa: codones (que son a la genética lo que los bits a la teoría de la información), traducción, transcripción, receptores, señal de comienzo, señal de parada, feedback, etcétera.

En un lenguaje, las palabras tienen sentido en el contexto en que se enuncian. Las moléculas biológicas tienen sentido sólo en el contexto en que ocurren. El enzima lo es del sustrato que activa o inhibe; el codón está en función del anticodón, el anticuerpo se fabrica contra un antígeno determinado, la célula crece, se multiplica y envejece de acuerdo con las instrucciones escritas en su núcleo. ¿Cómo se construye o surge esa información? En virtud de la estructura o secuencia y de la disposición espacial de las moléculas biológicas: los polímeros, en particular, los polipéptidos y los polinucleótidos. De la interacción entre polinuclétidos (ácidos nucleicos) y polipéptidos (proteínas) se desprende el proyecto y futuro desarrollo de la vida en todos sus niveles.

La interacción molecular es un tema clásico en física y en química, en el que se estudian la respuesta dieléctrica de sistemas moleculares y las fuerzas de interacción entre éstos, a largas y cortas distancias. Los polipéptidos y los polinucléotidos son macromoléculas que se rigen por las mismas leyes, con ciertas adaptaciones a su naturaleza específica. Especifidad que surge de la capacidad de autorreplicación de los polinucleótidos. Es decir, la información que el código conlleva no es estanca, sino que va creciendo en el tiempo y en el espacio con el paso de la evolución.

El comportamiento de los sistemas atómicos y moleculares puede preverse. Como predecirse la simetría de las partículas elementales aunque no esté comprobada en algún caso. Pero ¿puede predeterminarse cuál sería el comportamiento de las macromoléculas biológicas? Y aquí es donde la estadística, que le sirvió a Mendel para establecer su ley de la uniformidad de los descendientes y a Fisher, Haldane y Wright para dar un gran impulso a la matematización de la biología, viene en auxilio de los autores de esta monografía para crear sus modelos computacionales de simulación probable.

Para Fraga y coautores, los procesos de reconocimiento, de interacción, se fundan en consideraciones de tipo energético; la probabilidad de reconocimiento depende de los valores de las energías de estabilización de los distintos complejos portador-receptor. En su opinión, la teoría de la información molecular incorpora todas las formulaciones teóricas pertinentes de la interpretación cualitativa y de la predicción cuantitativa de los procesos de reconocimiento; la simulación por computador de interacciones y reacciones, con la ayuda de macroordenadores, es viable. Y la prueba de ello la encuentran en la bioingeniería capaz de diseñar nuevos enzimas, nuevos anticuerpos polipeptídicos, membranas inclusive.

No es la primera obra que acomete esta problemática. En España Martín Municio publicó en 1974 un espléndido volumen con el sugestivo título de *Interacciones Moleculares* y, más cercano en el tiempo, Eduardo Cadenas una completa monografía sobre el caso particular de los enzimas alostéricos (L. A.)

Bibliografía

Los lectores interesados en una mayor profundización de los temas expuestos pueden consultar los trabajos siguientes:

LA PROXIMA GENERACION DE ACELERADORES DE PARTICULAS

PROTON–PROTON INTERSECTING STORAGE ACCELERATOR FACILITY ISABELLE AT THE BROOKHAVEN NATIONAL LABORATORY. H. Hahn, M. Month y R. R. Rau, en *Reviews of Modern Physics*, vol. 49, n.° 3, págs. 625-679; julio, 1977.

DESIGN STUDY OF A 22 TO 130 GeV $e^+ e^-$ COLLIDING BEAM MACHINE (LEP). The LEP Study Group, European Organization for Nuclear Research, 1979.

FANTASIES OF FUTURE FERMILAB FACILITIES. R. R. Wilson, en *Reviews of Modern Physics*, vol. 51, n.° 2, págs. 259-273; abril, 1979.

BIOLOGIA CELULAR DEL ENVEJECIMIENTO HUMANO

CURRENT THEORIES OF BIOLOGICAL AGING. Leonard Hayflick, en *Federation Proceedings*, vol. 34, n.° 1, págs. 9-13; enero, 1975.

THE CELL BIOLOGY OF HUMAN AGING. Leonard Hayflick, en *The New England Journal of Medicine*, vol. 295, n.° 23, págs. 1302-1308; 2 de diciembre de 1976.

HANDBOOK OF THE BIOLOGY OF AGING. Dirigido por Caleb E. Finch y Leonard Hayflick. Van Nostrand Reinhold Co., 1977.

TIME, CELLS, AND AGING. Bernard L. Strehler. Academic Press, 1977.

THE BIOLOGY OF SENESCENCE. Alex Comfort. Elsevier North-Holland, Inc., 1979

RESERVAS URANIFERAS MUNDIALES

A SECOND IRON AGE AHEAD? Brian J. Skinner, en *American Scientist*, vol. 64, n.° 3, págs. 258-269; mayo-junio, 1976.

SHORT COURSE IN URANIUM DEPOSITS: THEIR MINERALOGY AND ORIGIN. Dirigido por M. M. Kimberley, Mineralogical Association of Canada. University of Toronto Press, 1978.

URANIUM DISTRIBUTION IN MINED DEPOSITS AND IN THE EARTH'S CRUST. Kenneth Deffeyes e Ian Mac-Gregor. U.S. Department of Energy, Grand Junction, Colo., agosto, 1978.

LOS SATELITES GALILEANOS DE JUPITER

MELTING OF IO BY DISSIPATION. S. J. Peale, P. Cassen y R. T. Reynolds, en *Science*, vol. 203, n.° 4383, págs. 892-894; 2 de marzo de 1979.

INFRARED OBSERVATIONS OF THE JOVIAN SYSTEM FROM VOYAGER 1. R. Hanel, B. Conrath, M. Fraser, V. Kunde, P. Lowman, W. Maguire, J. Pearl, J. Pirraglia, R. Samuelson, D. Gautier, P. Gierasch, S. Kumar y C. Ponnamperuma, en *Science*, vol. 204, n.° 4396, págs. 972-976; 1 de junio de 1979.

THE JUPITER SYSTEM THROUGH THE EYES OF VOYAGER 1. Bradford A. Smith, Laurence A. Soderblom, Torrence V. Johnson, Andrew P. Ingersoll, Stewart A. Collins, Eugene M. Shoemaker, G. E. Hunt, Harold Masursky, Michael H. Carr, Merton E. Davies, Allan F. Cook II, Joseph Boyce, G. Edward Danielson, Tobias Owen, Carl Sagan, Reta F. Beebe, Joseph Veverka, Robert G. Strom, John F. McCauley, David Morrison, Geoffrey A. Briggs y Verner E. Soumi, en *Science*, vol. 204, n.° 4396, págs. 951-972; 1 de junio de 1979.

VOYAGER 1 ENCOUNTER WITH THE JOVIAN SYSTEM. E. C. Stone y A. L. Lane, en *Science*, vol. 204, número 4396, páginas 945-948; primero de junio de 1979.

ILUSIONES GEOMETRICAS

A DEPTH PROCESSING THEORY OF THE POGGENDORFF ILLUSION. B. Gillam en *Perception and Psychophysics*, vol. 10, n.° 4A, págs. 211-216; 1971

THE PSYCHOLOGY OF VISUAL ILLUSION. J. O. Robinson. Hutchinson University Library, 1972.

ILLUSION IN NATURE AND ART. Dirigido por R. L. Gregory y E. H. Gombrich. Charles Scribner's Sons, 1973.

SEEING IS DECEIVING: THE PSYCHOLOGY OF VISUAL ILLUSIONS. Stanley Coren y Joan S. Girgus. Laurence Erlbaum Associates, 1978.

EL ORIGEN DEL MAIZ

THE ORIGIN OF CORN. Walton C. Galinat, en *Annual Review of Genetics*, vol. 5, págs. 447-478; 1971.

CORN: ITS ORIGIN, EVOLUTION AND IMPROVEMENT. Paul C. Mangelsdorf. Harvard University Press, 1974.

ZEA DIPLOPERENNIS (GRAMINEAE): A NEW TEOSINTE FROM MEXICO. Hugh H. Iltis, John F. Doebley, Rafael Guzmán y Batia Pazy, en *Science*, vol. 203, n.° 4376, págs. 186-188; 12 de enero de 1979.

LA RABIA

THE NATURAL HISTORY OF RABIES. Dirigido por George M. Baer. Academic Press, 1975.

SUCCESSFUL PROTECTION OF HUMANS EXPOSED TO RABIES INFECTION. Mahmoud Bahmanyar, Ahmad Fayaz, Shokrollah Nour-Salehi, Manouchehr Mohammadi e Hilary Koprowski, en *The Journal of the American Medical Association*, vol. 236, n.° 24, págs. 2751-2754; 13 de diciembre de 1976.

MONOCLONAL ANTIBODIES AGAINST RABIES VIRUS PRODUCED BY SOMATIC CELL HYBRIDIZATION: DETECTION OF ANTIGENIC VARIANTS. T. J. Wiktor y H. Koprowski, en *Proceedings of the National Academy of Sciences of the United States of America*, vol. 75, n.° 8, págs. 3938-3942; agosto, 1978.

ENZIMAS ALOSTERICOS

QUANTITATIVE ASPECTS OF ALLOSTERIC MECHANISMS. A. Levitzki. Springer-Verlag. 1978.

ENZIMAS ALOSTÉRICOS. COOPERATIVIDAD EN LA UNIÓN Y EN LA CATÁLISIS. E. Cadenas. H. Blume Ediciones. Madrid, 1978.

BIOLOGÍA MOLECULAR DE LA REGULACIÓN. E. Cadenas. Ediciones Omega S. A. *(En preparación.)* Barcelona.

JUEGOS MATEMATICOS

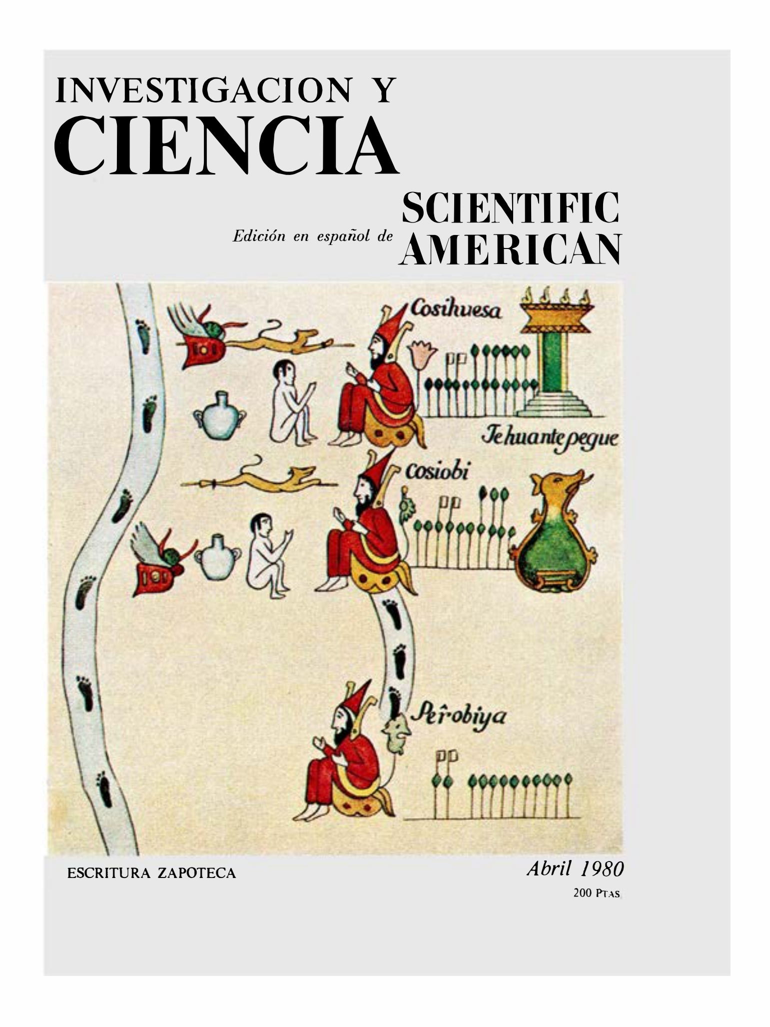
SOME STUDIES IN MACHINE LEARNING USING THE GAME OF CHECKERS. II: RECENT PROGRESS. A. L. Samuel, en *IBM Journal of Research and Development*, vol. 11, n.° 6, págs. 601-617; noviembre, 1967.

THE COMPLEXITY OF CHECKERS ON AN $N \times N$ BOARD-PRELIMINARY REPORT. A. S. Fraenkel, M. R. Garey, D. S. Johnson y Y. Yesha, en *19th Annual Symposium on the Foundations of Computer Science*. Institute of Electrical and Electronics Engineers, 1978.

TALLER Y LABORATORIO

ANALYTICAL ABSORPTION SPECTROSCOPY: ABSORPTIMETRY AND COLORIMETRY. Dirigido por M. G. Mellon. John Wiley & Sons, Inc., 1950.

Seguiremos explorando los campos del conocimiento



LA ASIGNACION DEL RADIOESPECTRO, por Charles Lee Jackson
El espectro electromagnético de 10 kilohertz a 300 gigahertz es un recurso natural de comunicación que debe asignarse de manera equitativa. Actualmente se están estudiando posibles reformas en el sistema de asignación.

ELEMENTOS GENETICOS MUDABLES, por Stanley N. Cohen y James A. Shapiro
No siguen las reglas de la recombinación genética ordinaria y unen segmentos de ADN que no están relacionados, transfiriendo grupos de genes de plásmidos, virus y cromosomas a las células vivas.

EL OBSERVATORIO EINSTEIN DE RAYOS X, por Riccardo Giacconi
Este satélite, que lleva un telescopio de rayos X de tamaño y sensibilidad sin precedentes, ha estado en el espacio durante más de un año. Ha proporcionado nuevas vistas espectaculares de fenómenos de altas energías en el universo.

LA PERCEPCION DEL CRECIMIENTO HUMANO, por James T. Todd, Leonard S. Mark, Robert E. Shaw y John B. Pittenger
¿Cómo cambia de forma la cabeza humana desde la infancia hasta la edad adulta? Los tests realizados muestran que el modo en que la mayoría de las personas perciben dicho crecimiento puede representarse mediante un determinado tipo de transformación geométrica.

PLANETAS CON ANILLOS, por Ignacio R. Ferrín V.
Hasta hace unos pocos años, sólo se conocían los anillos de Saturno. Desde entonces se han descubierto en torno a Urano y Júpiter. La estructura, tamaño de las partículas que los componen y constitución física de estos cuerpos siguen siendo objeto de discusión.

CANGREJOS DECORADORES, por Mary K. Wicksten
Muchas especies de arañas de mar pueden camuflarse empleando pautas de comportamiento mediante las que seleccionan materiales de su ambiente y los colocan formando guirnaldas sobre sus caparazones.

LA ESCRITURA ZAPOTECA, por Joyce Marcus
La escritura zapoteca figura entre las altas culturas de México anteriores a la conquista española. Sus jeroglifos, la mayoría de ellos esculpidos entre 500 a. C. y 700 d. C., nos han dejado constancia del apogeo y decadencia del estado zapoteca.

EL PARQUE DE YELLOWSTONE COMO VENTANA HACIA EL INTERIOR DE LA TIERRA, por Robert B. Smith y Robert L. Christiansen
Yellowstone es un "polvorín" en la corteza terrestre. Su gran actividad volcánica y tectónica le convierte en un lugar único para el estudio de procesos que se originan en el interior de la Tierra.

INVESTIGACION Y
CIENCIA